(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
13 May 2004 (13.05.2004)

PCT

(10) International Publication Number
WO 2004/039311 A2

(51) International Patent Classification[7]: A61K

(21) International Application Number:
PCT/US2003/015722

(22) International Filing Date: 15 May 2003 (15.05.2003)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/381,027	15 May 2002 (15.05.2002)	US
60/382,702	21 May 2002 (21.05.2002)	US
60/451,146	28 February 2003 (28.02.2003)	US

(71) Applicant *(for all designated States except US)*: **UNIVERSITY OF UTAH RESEARCH FOUNDATION** [US/US]; 615 Arapeen Drive, Suite 110, Salt Lake City, UT 84108 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **SUNDQUIST, Wesley, I.** [US/US]; 299 Fairfax Circle, Salt Lake City, UT 84103 (US). **GARRUS, Jennifer, E.** [US/US]; 2686 Preston Street, Salt Lake City, UT 84105 (US). **VON SCHWEDLER, Uta, K.** [US/US]; 1123 South Lake Street, Salt Lake City, UT 84105 (US).

(74) Agents: **HUIZENGA, David, E.** et al.; Needle & Rosenberg, P.C., The Candler Building, 127 Peachtree Street, Atlanta, GA 30303-1811 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 2004/039311 A2

(54) Title: HUMAN VPS4A AND VPS4B FUNCTIONS IN VIRAL BUDDING

(57) Abstract: Disclosed are compositions and methods related to retroviral budding.

HUMAN VPS4A AND VPS4B FUNCTIONS IN VIRAL BUDDING

1. This application claims the benefit of and priority to United States Provisional Application No. 60/381,027 filed on May 15, 2002 for Human Vps4A and Vps4B Functions in Viral Budding, United States Provisional Application No. 60/382,702 filed on May 21, 2002 for Human Vps4A and Vps4B Functions in Viral Budding, and United States Provisional Application No. 60/451,146 filed on February 28, 2003 for Human Vps4A and Vps4B Functions in Viral Budding. These applications are specifically herein incorporated by reference in their entireties.

I. ACKNOWLEDMENTS

2. This invention was made with government support under federal grants AI51174 awarded by the NIH. The Government has certain rights to this invention.

II. BACKGROUND OF THE INVENTION

3. Human Tsg101 plays key roles in HIV budding and in cellular vacuolar protein sorting (Vps). In performing these functions, Tsg101 binds both ubiquitin (Ub) and the PTAP tetrapeptide "late domain" motif located within the viral Gag protein. These interactions are mediated by the N-terminal domain of Tsg101, which belongs to the catalytically inactive ubiquitin E2 variant (UEV) family.

4. Human Tsg101 has recently been identified as the functional receptor required for budding of the enveloped human immunodeficiency (HIV) and Ebola viruses (VerPlank et al., 2001; Garrus et al., 2001; Martin-Serrano et al., 2001, all of which are herein incorporated by reference at least for material related to TSG101). Tsg101 is recruited to the sites of virus budding by binding to a PTAP tetrapeptide motif (the "late domain") located within the p6 region of HIV Gag and the Ebola Vp40 matrix structural proteins. Once there, Tsg101 appears to recruit other cellular factors that help to complete the budding process (Garrus et al., 2001; Martin-Serrano et al., 2001). In the absence of Tsg101, the final membrane fission step(s) fail and HIV release is arrested at a very late stage, in which the assembled viral particles remain attached to the plasma membrane (and to one another) via thin membrane stalks (Garrus et al., 2001).

5. Disclosed herein are relationships and interactions related to TSG101 and Vps4A and 4B that are integral for HIV budding.

III. SUMMARY OF THE INVENTION

6. In accordance with the purposes of this invention, as embodied and broadly described herein, this invention, in one aspect, relates to the HIV budding.

7. Additional advantages of the invention will be set forth in part in the description which follows, and in part will be obvious from the description, or may be learned by practice of the invention. The advantages of the invention will be realized and attained by means of the elements and combinations particularly pointed out in the appended claims. It is to be understood that both the foregoing general description and the following detailed description are exemplary and explanatory only and are not restrictive of the invention, as claimed.

IV. BRIEF DESCRIPTION OF THE DRAWINGS

8. The accompanying drawings, which are incorporated in and constitute a part of this specification, illustrate several embodiments of the invention and together with the description, serve to explain the principles of the invention.

9. Figure 1 shows human Tsg101 Binds HIV-1 p6 (A) HIV-1 Gag domain structure and dissociation constants for pure Tsg101 UEV protein binding to immobilized HIV-1 GST-p6 and GST-p6-Ub proteins. The standard deviation in the dissociation constant for the full length, wild type p6 protein was derived from 16 independent measurements; estimated errors in the final significant digits of the other dissociation constants (parentheses) were derived from statistical fits of individual isotherms to 1:1 binding models. (**B**) SDS-PAGE analysis showing induced expression in *E. coli* and fully purified Tsg101 UEV. (**C**) Surface plasmon resonance biosensor analysis of the p6/Tsg 101 interaction. Tsg101 UEV was injected in triplicate at concentrations of 0, 0.9, 0.27, 0.82, 2.5, 7.4, 22, 67, and 200 μM over GST-p6 captured on an anti-GST surface. Data were double referenced to correct for bulk refractive index changes and any nonspecific binding (Myszka, 1999). (**D**) Representative binding isotherms for Tsg101 UEV interacting with a series of different protein surfaces. (**E**) Epitope mapping of the Tsg101 UEV binding site on HIV-1 p6. The graph shows the changes in free energy of Tsg101 UEV binding (ΔΔG) caused by single alanine substitutions in HIV-1 p6 residues 4-13 (dark bars) or by the P7L, T8I, A9R, and P10L mutations (light bars). Positive values indicate that the mutation reduces binding affinity, and error bars show estimated standard deviations in single binding isotherms. Broken lines above the P10L mutant denote a lack of detectable Tsg101 UEV binding (see also Fig. 7D).

10. Figure 2 shows Tsg101 is Required for Release of HIV-1 from 293T Cells (**A**) Anti-Tsg101 Western blot showing depletion of Tsg101 protein from 293T cells by siRNA and restoration of Tsg101 protein using siRNA-resistant expression constructs. Cells were co-transfected twice (2 mL cultures, 6 well plates) with plasmid DNA and with 50 nM siRNA or $siRNA_{INV}$ at 24h intervals (except mock, lane 1, which received no RNA or DNA). The first siRNA/$siRNA_{INV}$ co-transfection was with 2μg or 4μg of a Tsg-FLAG (siRNA-sensitive) or Tsg*-FLAG (siRNA-resistant) expression construct. The second siRNA/$siRNA_{INV}$ co-transfection was with 0.5μg HIV-1 R9 DNA. In each case, total transfected DNA was normalized to 4.5 μg with pIRES2-EGFP. Cytoplasms and supernatants were analyzed 48h after the first transfection. (**B**) Anti-FLAG Western blot showing that the Tsg101*-FLAG construct is resistant to siRNA. Samples were prepared as in (A). (C) Anti-MA and anti-CA Western blot showing that depletion of Tsg101 inhibits virus release. From cells transfected as in (A), viral particles were pelleted from the supernatant of the transfected cells. Virus release in untreated cells was similar to that in the $siRNA_{INV}$-treated cells (not shown). (D) Anti-MA and anti-CA Western blot showing that cytoplasmic levels of HIV-1 Gag, CA and MA proteins are not affected by Tsg101 depletion. Note that levels of MA and CA in cellular membrane fractions increased significantly when virus release was blocked (not shown). (E) Viral replication assays showing that release of infectious virus particles is inhibited when Tsg101 is

depleted and rescued when Tsg101 is restored. Titers of infectious HIV-1 particles released into the supernatant (shown) were quantitated in single cycle MAGIC assays.

11. Figure 3 shows the depletion of Tsg101 Does Not Affect MLV Release or Infectivity. (**A**) Depletion of Tsg101 does not significantly inhibit MLV release from 293T cells. Upper panel: Anti-Tsg101 and anti-14-3-3β Western blot showing specific depletion of Tsg101 from 293T cells by siRNA. Depletion experiments were performed as described in Fig. 8A, except that the second siRNA co-transfection was with 1.5 μg pCLeco and 0.25 μg pCLlacZ MLV vector DNA. Levels of 14-3-3β protein are shown as loading controls, and a pCLeco vector expressing PPPY⁻ mutant Gag is shown as a positive control for arrested budding (lane 7). Middle panel: Anti-MLV CA Western blot showing that virus release is not significantly reduced by Tsg101 depletion. Lower panel: Anti-CA Western blot showing cytoplasmic levels of MLV Gag. (**B**) Depletion of Tsg101 does not significantly inhibit MLV infectivity. Infectious MLV particles released into the supernatant were quantitated in single cycle infectivity assays and normalized to the appropriate controls. In repetitions of these experiments, siRNA treatment reduced particle release and infectivity slightly (up to two-fold) relative to the $siRNA_{INV}$ control.

12. Figure 4 shows a Mutant Vps4 Proteins Dominantly Inhibit HIV-1 Release and Infectivity (**A**) Mutant Vps4 proteins dominantly inhibit HIV-1 release. 293T cells were transfected with 0.5 μg R9 plasmid encoding wt HIV-1 (lanes 2, 4-10) or a p6 PTAP⁻ mutant HIV-1 (lane 3). Cells were co-transfected with pEGFP (1 μg, lane 2), GFP-Vps4 wt (1 μg, lane 4), GFP-$Vps4_{K173Q}$ (0.25, 0.5 or 1 μg; lanes 5-7), or GFP-$Vps4_{E228Q}$ (0.25, 0.5 or 1 μg; lanes 8-10). Total DNA was normalized to 1.5 μg with pEGFP. Pelleted virus (above) and cytoplasmic extracts (labeled Cell, below) were harvested after 24 h, and levels of expressed CA, MA, and GFP-Vps4 proteins analyzed by Western blotting. (**B**) Mutant Vps4 proteins dominantly inhibit HIV-1 infectivity. Levels of infectious HIV-1 particles released into the supernatant were quantitated using single cycle MAGIC assays. Samples were from (**A**). (**C**) Mutant Vps4 proteins dominantly inhibit release of Gag-GFP. 293T cells were transfected with 1 μg vector expressing wt Gag-GFP (lanes 2, 4-6) or a p6 PTAP⁻ mutant Gag-GFP (lane 3). Samples were co-transfected with 1 μg vector expressing GFP only (lane 2), wt GFP-Vps4 (lane 4), GFP-$Vps4_{K173Q}$ (lane 5), or GFP-$Vps4_{E228Q}$ (lane 6). Pelleted virus (above) and cytoplasmic extracts (below) were harvested after 24 h, and Gag and GFP-Vps4 expression were analyzed by Western blotting with anti-MA, anti-CA (top two panels) and anti-GFP antibodies (bottom panel). (**D**) PTAP- mutation and dominant negative Vps4 proteins inhibit release of infectious HIV-1 from MT4 T cells. Cells were transfected (24 well plates) with 1 μg HIV-1 R9 DNA and 1 μg Vps4 DNA or 1 μg control DNA (pEGFP). Levels of infectious HIV-1 particles released into the supernatant were quantitated using single cycle MAGIC assays 72 h post transfection.

13. Figure 5 shows Mutant Vps4 Proteins Dominantly Inhibit MLV Release and Infectivity. (**A**) 293T cells were co-transfected with 1.5 μg pCLeco and 0.25 μg pCLlacZ MLV vectors encoding wt (lanes 2,4-6) or PPPY⁻ mutant MLV Gag proteins (lane 3). Co-transfections were with 1 μg of the following vectors: pEGFP expressing GFP only (lanes 2, 3), wt GFP-Vps4 (lane 4), GFP-$Vps4_{K173Q}$ (lane 5), or GFP-$Vps4_{E228Q}$ (lane 6). Pelleted virus (upper panel) and cytoplasmic extracts (lower panels) were

harvested after 24 h, and levels of expressed MLV CA, Gag, and GFP-Vps4 proteins were analyzed by Western blotting with anti-MLV CA and anti-GFP antibodies as indicated. (**B**) Mutant Vps4 proteins dominantly inhibit MLV replication. Samples were from (A) and MLV vector titers were quantitated as described in Experimental Methods.

14. Figure 6. Tsg101 Depletion and Vps4 Mutants Block Virus Budding at a Late Stage. (**A**) Representative thin section electron micrographs showing the effects on HIV-1 budding of mutating the p6 PTAP sequence (left), depleting cellular Tsg101 (center), or overexpressing the dominant negative $Vps4_{E228Q}$ protein (right). In each case, virus budding arrested at a late stage, with immature particles remaining connected to the plasma membrane via membrane stalks or to other budding particles to form "clusters" of interconnected particles. A subset of the cells transfected with the siRNA and mutant Vps4 constructs exhibited what appeared to be Class E compartments in the plane of section, and a full description will be published elsewhere. Scale bars are all 100 nm. (**B**) Representative thin section electron micrographs showing the effects on MLV budding of mutating the p12 PPPY motif (left) or overexpressing the dominant negative $Vps4_{E228Q}$ protein (right). These MLV late domain phenotypes are similar to those described above for HIV-1.

15. Figure 7 shows that Vps4b mutants block budding of HIV GAG. Figure 7A shows a Western blot of proteins isolated from 293T cells co-transfected with the indicated constructs. Figure 7B shows immunofluorescence of the GsRed and GFP labeled proteins.

16. Figure 8 shows that Vps4b blocks virus budding and infectivity. Figure 8A shows a Western blot of proteins isolated from cells co-transfected with HIV virus and the indicated constructs. Figure 8B shows the results of a MAGIC assay of the co-transfected cells plotting the various constructs vs the infectivity titer.

17. Figure 9 shows that BC2 overexpression blocks HIV budding. Figure 9B shows a Western blot of proteins isolated from cells after a GST pulldown experiment of cells cotransfected with the indicated constructs. Figure 9C shows a Western blot of proteins isolated from cells co-transfected with the indicated constructs and HIV virus. Figure 9D shows the results of a MAGIC infectivity assay on cells co-transfected with the indicated constructs.

18. Figure 10 shows a schematic of molecules involved in viral budding.

V. DETAILED DESCRIPTION

19. The present invention may be understood more readily by reference to the following detailed description of preferred embodiments of the invention and the Examples included therein and to the Figures and their previous and following description.

20. Before the present compounds, compositions, articles, devices, and/or methods are disclosed and described, it is to be understood that this invention is not limited to specific synthetic methods, specific recombinant biotechnology methods unless otherwise specified, or to particular reagents unless otherwise specified, as such may, of course, vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only and is not intended to be limiting.

A. Definitions

21. As used in the specification and the appended claims, the singular forms "a," "an" and "the" include plural referents unless the context clearly dictates otherwise. Thus, for example, reference to "a pharmaceutical carrier" includes mixtures of two or more such carriers, and the like.

22. Ranges may be expressed herein as from "about" one particular value, and/or to "about" another particular value. When such a range is expressed, another embodiment includes from the one particular value and/or to the other particular value. Similarly, when values are expressed as approximations, by use of the antecedent "about," it will be understood that the particular value forms another embodiment. It will be further understood that the endpoints of each of the ranges are significant both in relation to the other endpoint, and independently of the other endpoint. It is also understood that there are a number of values disclosed herein, and that each value is also herein disclosed as "about" that particular value in addition to the value itself. For example, if the value "10" is disclosed, then "about 10" is also disclosed. It is also understood that when a value is disclosed that "less than or equal to" the value, "greater than or equal to the value" and possible ranges between values are also disclosed, as appropriately understood by the skilled artisan. For example, if the value "10" is disclosed then "less than or equal to 10"as well as "greater than or equal to 10" is also disclosed.

23. Throughout this application various assays and methods are disclosed for determining inhibition or inhibitors. It is understood that these methods are also disclosed as being practiced, for example, by modulating or modulators, or enhancing or enhancers (molecules that increase a certain activity). In other words, the particular activity assayed for, inhibition, modulation, or enhancing can all be equally used.

24. In this specification and in the claims which follow, reference will be made to a number of terms which shall be defined to have the following meanings:

25. "Optional" or "optionally" means that the subsequently described event or circumstance may or may not occur, and that the description includes instances where said event or circumstance occurs and instances where it does not.

26. Fragments. It is understood that where peptides or nucleic acid molecules are described, fragments of these molecules are also disclosed, unless specifically indicated otherwise. For example, if a Bc2 is disclosed in a particular method or assay, it is understood that fragments of Bc2 are also disclosed therein, unless specifically indicated to the contrary.

27. Throughout this application, various publications are referenced. The disclosures of these publications in their entireties are hereby incorporated by reference into this application in order to more fully describe the state of the art to which this invention pertains. The references disclosed are also individually and specifically incorporated by reference herein for the material contained in them that is discussed in the sentence in which the reference is relied upon.

28. It will be apparent to those skilled in the art that various modifications and variations can be made in the present invention without departing from the scope or spirit of the invention. Other

embodiments of the invention will be apparent to those skilled in the art from consideration of the specification and practice of the invention disclosed herein. It is intended that the specification and examples be considered as exemplary only, with a true scope and spirit of the invention being indicated by the following claims.

29. Disclosed are the components to be used to prepare the disclosed compositions as well as the compositions themselves to be used within the methods disclosed herein. These and other materials are disclosed herein, and it is understood that when combinations, subsets, interactions, groups, etc. of these materials are disclosed that while specific reference of each various individual and collective combinations and permutation of these compounds may not be explicitly disclosed, each is specifically contemplated and described herein. For example, if a particular TSG 101 variant is disclosed and discussed and a number of modifications that can be made to a number of molecules including the TSG 101 variant are discussed, specifically contemplated is each and every combination and permutation of the TSG 101 variant and the modifications that are possible unless specifically indicated to the contrary. Thus, if a class of molecules A, B, and C are disclosed as well as a class of molecules D, E, and F and an example of a combination molecule, A-D is disclosed, then even if each is not individually recited each is individually and collectively contemplated meaning combinations, A-E, A-F, B-D, B-E, B-F, C-D, C-E, and C-F are considered disclosed. Likewise, any subset or combination of these is also disclosed. Thus, for example, the sub-group of A-E, B-F, and C-E would be considered disclosed. This concept applies to all aspects of this application including, but not limited to, steps in methods of making and using the disclosed compositions. Thus, if there are a variety of additional steps that can be performed it is understood that each of these additional steps can be performed with any specific embodiment or combination of embodiments of the disclosed methods.

B. Compositions and Methods

30. Disclosed herein are mechanisms and pathways supported by data which are consistent with an ESCRT-1 complex selecting HIV Gag and other viral and non-viral PTAP proteins as cargo. TSG101 binds HIV Gag-p6 and VPS28 during this selection. The VPS28 binding to the TSG101 S-box is tight (8nM) and long-lived (t1/2>1h), and VPS28 is unstable when not bound to the TSG101 S-box. It is also disclosed herein that VPS28 (and/or the S-box) are required for HIV budding. The data disclosed herein indicate that other Vps proteins are involved downstream of these interactions. ATPase-defective mutants of VPS4a and 4b disrupt virus budding. Furthermore, disclosed herein, BC2/CHMP2 binds to VPS4a/b and BC2/CHMP2 overexpression (or fusion to DsRed) inhibits HIV budding. Disclosed are compositions and methods for addressing the effect of Bc2 on viral budding.

31. Tsg101 binds to other viral and non-viral PTAP sequences. It also binds to human Vps28, tightly and with a long half life, and this complex is required for ESCRT-1 function in HIV budding.

32. Disclosed are interactions between TSG101 and any other viral genes with a PTAP sequence, such as Ebola Vp40, Blue tongue virus, and PPPY viruses.

33. The binding sites of TSG101, Vps4A and Vps4B, Bc2, Vps28, Vps37, and GAG molecules, as disclosed herein set forth targets for interaction with ligands which will be able to bind or modulate, such as inhibit, the disclosed interactions.

34. A compound that is identified or designed as a result of any of the disclosed methods can be obtained (or synthesized) and tested for its biological activity, e.g., inhibition of TSG101-GAG interaction activity.

1. TSG101-Ub interactions and effects

35. Disclosed are methods for reducing interactions between TSG101and Ub, comprising incubating an inhibitor of the interaction between TSG101 and Ub. Also disclosed are methods for inhibiting retroviral budding comprising administering an inhibitor of the interaction between TSG101 and Ub.

36. Disclosed are methods of treating a subject comprising administering to the subject an inhibitor of HIV budding, wherein the inhibitor reduces the interaction between TSG101 and Ub, and wherein the subject is in need of such treatment.

37. Also disclosed are methods, wherein the inhibitor disrupts an interaction between TSG101 and the Ub, and/or wherein the inhibitor interacts with the UEV domain of TSG101.

38. Disclosed are methods of identifying an inhibitor of an interaction between TSG101 and Ub comprising incubating a library of molecules with TSG101 forming a mixture, and identifying the molecules that disrupt the interaction between TSG101 and Ub, wherein the interaction disrupted comprises an interaction between the Ub and an amino acid of TSG101.

39. Also disclosed are methods, wherein the step of isolating comprises incubating the mixture with a molecule comprising Ub.

40. Disclosed are methods of identifying an inhibitor of an interaction between TSG101 and Ub comprising incubating a library of molecules with Ub forming a mixture, and identifying the molecules that disrupt the interaction between Ub and TSG101, wherein the interaction disrupted comprises an interaction between the Ub and an amino acid of TSG101.

41. Also disclosed are the methods, wherein the step of isolating comprises incubating the mixture with a molecule comprising a TSG 101 UEV domain.

42. Also disclosed are compositions produced by any of the processes as disclosed herein, as well as compositions capable of being identified by the processes disclosed herein.

43. It is understood that the disclosed methods and compositions are active for retroviruses dependent on GAG, such as HIV and Ebola.

44. It is also understood that the disclosed methods and compositions can be further combined, with for example, PTAP domain addition, or any other protein or composition disclosed herein. Disclosed herein, Ub and PTAP domains bind TSG101 such that the binding of the two in the presence of each other is tight then the binding of either alone. Thus, disclosed are methods and assays where, for

example, both PTAP domains and Ub are incubated with target molecules or used in systems to assays for molecules that inhibit retroviral budding.

45. Disclosed are methods of manufacturing a composition for inhibiting the interaction between TSG101 and Ub comprising synthesizing the inhibitors as disclosed herein.

46. Also disclosed are methods that include mixing a pharmaceutical carrier with the inhibitor as disclosed herein, and produced by any of the disclosed methods.

47. Disclosed are methods of identifying inhibitors of TSG101-Ub interaction comprising, a) administering a composition to a system, wherein the system supports TSG101-Ub interaction, b) assaying the effect of the composition on the amount of TSG101-Ub is in the system, and c) selecting a composition which causes a decrease in the amount of TSG101-Ub present in the system relative to the system without the addition of the composition.

48. Also disclosed are methods of identifying inhibitors of HIV budding comprising, a) administering a composition to a system, wherein the system supports HIV budding via a TSG101-Ub interaction, b) assaying the effect of the composition on the amount of HIV budding in the system, and c) selecting a composition which causes a decrease in the amount of HIV budding present in the system because of an inhibition of the TSG101-Ub interaction relative to the system without the addition of the composition.

49. Disclosed are methods of inhibiting HIV budding comprising administering a composition, wherein the composition prevents HIV budding, wherein the composition is defined as a composition capable of being identified by administering the composition to a system, wherein the system supports HIV-budding via a TSG101-Ub interaction, assaying the effect of the composition on the amount of HIV budding in the system, and selecting a composition which causes a decrease in the amount of HIV budding present in the system because of an inhibition of the TSG101-Ub interaction relative to the system without the addition of the composition.

50. Also disclosed are methods of inhibiting HIV budding comprising administering a composition that reduces an interaction between TSG101 and Ub.

51. Disclosed are methods of making a composition capable of inhibiting HIV budding comprising admixing a compound with a pharmaceutically acceptable carrier, wherein the compound is identified by administering the compound to a system, wherein the system supports HIV budding via a TSG101-Ub interaction, assaying the effect of the compound on the amount of HIV budding in the system, and selecting a compound which causes a decrease in the amount of HIV budding in the system because of an inhibition of the TSG101-Ub interaction, relative to the system without the addition of the compound.

52. Disclosed are methods of manufacturing an inhibitor to HIV budding comprising, a) administering a composition to a system, wherein the system supports HIV budding via a TSG101-Ub interaction, b) assaying the effect of the composition on the amount of HIV budding in the system, c) selecting a composition which causes a decrease in the amount of HIV budding present in the system

because of an inhibition of the TSG101-Ub interaction, relative to the system with the addition of the composition, and d) synthesizing the composition.

53. Also disclosed are methods comprising the step of admixing the composition with a pharmaceutical carrier.

54. Disclosed are methods of identifying an inhibitor of an interaction between TSG101 and Ub comprising a) administering a composition to a system, wherein the system comprises TSG101, b) assaying the effect of the composition on a TSG101-Ub interaction, and c) selecting a composition which inhibits a TSG101-Ub interaction.

55. Disclosed are cells that further comprising an inhibitor of a TSG101-Ub interaction.

2. Vps4a and Vps4b ATPase activity and effects

56. Disclosed are methods of identifying inhibitors of Vps4A or 4B ATPase activity comprising, a) administering a composition to a system, wherein the system supports Vps4A or 4B ATPase activity, b) assaying the effect of the composition on the amount of Vps4A or 4B ATPase activity in the system, and c) selecting a composition which causes a decrease in the amount of Vps4A or 4B ATPase activity present in the system relative to the system without the addition of the composition. Further disclosed are systems wherein the system comprises retroviral budding activity and the Vps4A or 4B ATPase activity is determined by assaying the retroviral budding activity.

57. Also disclosed are methods of identifying inhibitors of HIV budding comprising, a) administering a composition to a system, wherein the system supports HIV budding via a Vps4A or 4B ATPase activity, b) assaying the effect of the composition on the amount of HIV budding in the system, and c) selecting a composition which causes a decrease in the amount of HIV budding present in the system because of an inhibition of the Vps4A or 4B ATPase activity relative to the system without the addition of the composition.

58. Disclosed are methods of inhibiting HIV budding comprising administering a composition, wherein the composition prevents HIV budding, wherein the composition is defined as a composition capable of being identified by administering the composition to a system, wherein the system supports HIV-budding via a Vps4A or 4B ATPase activity interaction, assaying the effect of the composition on the amount of HIV budding in the system, and selecting a composition which causes a decrease in the amount of HIV budding present in the system because of an inhibition of the Vps4A or 4B ATPase activity relative to the system without the addition of the composition.

59. Also disclosed are methods of inhibiting HIV budding comprising administering a composition that reduces Vps4A or 4B ATPase activity.

60. Disclosed are methods of making a composition capable of inhibiting HIV budding comprising admixing a compound with a pharmaceutically acceptable carrier, wherein the compound is identified by administering the compound to a system, wherein the system supports HIV budding via a Vps4A or 4B ATPase activity, assaying the effect of the compound on the amount of HIV budding in the system, and selecting a compound which causes a decrease in the amount of HIV budding in the system

because of an inhibition of the Vps4A or 4B ATPase activity, relative to the system without the addition of the compound.

61. Disclosed are methods of manufacturing an inhibitor to HIV budding comprising, a) administering a composition to a system, wherein the system supports HIV budding via a Vps4A or 4B ATPase activity, b) assaying the effect of the composition on the amount of HIV budding in the system, c) selecting a composition which causes a decrease in the amount of HIV budding present in the system because of an inhibition of the Vps4A or 4B ATPase activity, relative to the system with the addition of the composition, and d) synthesizing the composition.

62. Also disclosed are methods comprising the step of admixing the composition with a pharmaceutical carrier.

63. Disclosed are methods of identifying an inhibitor of Vps4A or 4B ATPase activity comprising a) administering a composition to a system, wherein the system comprises Vps4A or 4B ATPase activity, b) assaying the effect of the composition on a Vps4A or 4B ATPase activity, and c) selecting a composition which inhibits a Vps4A or 4B ATPase activity.

64. Disclosed are cells comprising, a) a regulatable nucleic acid comprising sequence encoding a Vps4A or 4B gene and b) a nucleic acid comprising sequence encoding a budding retrovirus as well as cells further comprising a regulatable nucleic acid comprising sequence encoding a TSG101 gene.

65. Disclosed are cells that further comprising an inhibitor of a Vps4A or 4B ATPase activity.

3. Vps4a-Vps4b interactions and effects

66. Disclosed are methods for reducing interactions between Vps4a and Vps4b, comprising incubating an inhibitor of the interaction between Vps4a and Vps4b. Also disclosed are methods for inhibiting retroviral budding comprising administering an inhibitor of the interaction between Vps4a and Vps4b.

67. Disclosed are methods of treating a subject comprising administering to the subject an inhibitor of HIV budding, wherein the inhibitor reduces the interaction between Vps4a and Vps4b, and wherein the subject is in need of such treatment.

68. Also disclosed are methods, wherein the inhibitor disrupts an interaction between Vps4a and the Vps4b.

69. Disclosed are methods of identifying an inhibitor of an interaction between Vps4a and Vps4b comprising incubating a library of molecules with Vps4a forming a mixture, and identifying the molecules that disrupt the interaction between Vps4a and Vps4b, wherein the interaction disrupted comprises an interaction between the Vps4b and an amino acid of Vps4a.

70. Also disclosed are methods, wherein the step of isolating comprises incubating the mixture with a molecule comprising Vps4a and/or Vps4b.

71. Disclosed are methods of identifying an inhibitor of an interaction between Vps4a and Vps4b comprising incubating a library of molecules with Vps4b forming a mixture, and identifying the

molecules that disrupt the interaction between Vps4a and Vps4b, wherein the interaction disrupted comprises an interaction between the Vps4a and an amino acid of Vps4b.

72. Also disclosed are compositions produced by any of the processes as disclosed herein, as well as compositions capable of being identified by the processes disclosed herein.

73. It is understood that the disclosed methods and compositions are active for retroviruses dependent on GAG, such as HIV and Ebola.

74. It is also understood that the disclosed methods and compositions can be further combined, with for example, any other protein or composition disclosed herein, such as the PTAP domain. Thus, disclosed are methods and assays where, for example, Vps4a and Vps4b and Ub are incubated with target molecules or used in systems to assays for molecules that inhibit retroviral budding.

75. Disclosed are methods of manufacturing a composition for inhibiting the interaction between Vps4a and Vps4b comprising synthesizing the inhibitors as disclosed herein.

76. Also disclosed are methods that include mixing a pharmaceutical carrier with the inhibitor as disclosed herein, and produced by any of the disclosed methods.

77. Disclosed are methods of identifying inhibitors of Vps4a-Vps4b interaction comprising, a) administering a composition to a system, wherein the system supports Vps4a-Vps4b interaction, b) assaying the effect of the composition on the amount of Vps4a-Vps4b is in the system, and c) selecting a composition which causes a decrease in the amount of Vps4a-Vps4b present in the system relative to the system without the addition of the composition.

78. Also disclosed are methods of identifying inhibitors of HIV budding comprising, a) administering a composition to a system, wherein the system supports HIV budding via a Vps4a-Vps4b interaction, b) assaying the effect of the composition on the amount of HIV budding in the system, and c) selecting a composition which causes a decrease in the amount of HIV budding present in the system because of an inhibition of the Vps4a-Vps4b interaction relative to the system without the addition of the composition.

79. Disclosed are methods of inhibiting HIV budding comprising administering a composition, wherein the composition prevents HIV budding, wherein the composition is defined as a composition capable of being identified by administering the composition to a system, wherein the system supports HIV-budding via a Vps4a-Vps4b interaction, assaying the effect of the composition on the amount of HIV budding in the system, and selecting a composition which causes a decrease in the amount of HIV budding present in the system because of an inhibition of the Vps4a-Vps4b interaction relative to the system without the addition of the composition.

80. Also disclosed are methods of inhibiting HIV budding comprising administering a composition that reduces an interaction between Vps4a and Vps4b.

81. Disclosed are methods of making a composition capable of inhibiting HIV budding comprising admixing a compound with a pharmaceutically acceptable carrier, wherein the compound is identified by administering the compound to a system, wherein the system supports HIV budding via a

Vps4a-Vps4b interaction, assaying the effect of the compound on the amount of HIV budding in the system, and selecting a compound which causes a decrease in the amount of HIV budding in the system because of an inhibition of the Vps4a-Vps4b interaction, relative to the system without the addition of the compound.

82. Disclosed are methods of manufacturing an inhibitor to HIV budding comprising, a) administering a composition to a system, wherein the system supports HIV budding via a Vps4a-Vps4b interaction, b) assaying the effect of the composition on the amount of HIV budding in the system, c) selecting a composition which causes a decrease in the amount of HIV budding present in the system because of an inhibition of the Vps4a-Vps4b interaction, relative to the system with the addition of the composition, and d) synthesizing the composition.

83. Also disclosed are methods comprising the step of admixing the composition with a pharmaceutical carrier.

84. Disclosed are methods of identifying an inhibitor of an interaction between Vps4a and Vps4b comprising a) administering a composition to a system, wherein the system comprises Vps4a, b) assaying the effect of the composition on a Vps4a-Vps4b interaction, and c) selecting a composition which inhibits a Vps4a-Vps4b interaction.

85. Disclosed are methods of identifying an inhibitor of an interaction between Vps4a and Vps4b comprising a) administering a composition to a system, wherein the system comprises Vps4b, b) assaying the effect of the composition on a Vps4a-Vps4b interaction, and c) selecting a composition which inhibits a Vps4a-Vps4b interaction.

86. Disclosed are cells that further comprising an inhibitor of a Vps4a-Vps4b interaction.

4. Vps4a-Bc2 interactions and effects

87. Disclosed are methods for reducing interactions between Vps4a and Bc2, comprising incubating an inhibitor of the interaction between Vps4a and Bc2. Also disclosed are methods for inhibiting retroviral budding comprising administering an inhibitor of the interaction between Vps4a and Bc2 or administration of Bc2, such as excess Bc2.

88. Disclosed are methods of treating a subject comprising administering to the subject an inhibitor of HIV budding, wherein the inhibitor reduces the interaction between Vps4a and Bc2, and wherein the subject is in need of such treatment, and wherein the inhibitor is Bc2 or analog.

89. Also disclosed are methods, wherein the inhibitor disrupts an interaction between Vps4a and the Bc2.

90. Disclosed are methods of identifying an inhibitor of an interaction between Vps4a and Bc2 comprising incubating a library of molecules with Vps4a forming a mixture, and identifying the molecules that disrupt the interaction between Vps4a and Bc2, wherein the interaction disrupted comprises an interaction between the Bc2 and an amino acid of Vps4a.

91. Also disclosed are methods, wherein the step of isolating comprises incubating the mixture with a molecule comprising Bc2 or fragment thereof.

92. Disclosed are methods of identifying an inhibitor of an interaction between Vps4a and Bc2 comprising incubating a library of molecules with Bc2 forming a mixture, and identifying the molecules that disrupt the interaction between Vps4a and Bc2, wherein the interaction disrupted comprises an interaction between the Bc2 and an amino acid of Vps4a.

93. Also disclosed are methods, wherein the step of isolating comprises incubating the mixture with a molecule comprising Vps4a or fragment thereof.

94. Also disclosed are compositions produced by any of the processes as disclosed herein, as well as compositions capable of being identified by the processes disclosed herein.

95. It is understood that the disclosed methods and compositions are active for retroviruses dependent on GAG, such as HIV and Ebola.

96. Disclosed are methods of manufacturing a composition for inhibiting the interaction between Vps4a and Bc2 comprising synthesizing the inhibitors as disclosed herein.

97. Also disclosed are methods that include mixing a pharmaceutical carrier with the inhibitor as disclosed herein, and produced by any of the disclosed methods.

98. Disclosed are methods of identifying inhibitors of Vps4a and Bc2 interaction comprising, a) administering a composition to a system, wherein the system supports Vps4a-Bc2 interaction, b) assaying the effect of the composition on the amount of Vps4a-Bc2 is in the system, and c) selecting a composition which causes a decrease in the amount of Vps4a-Bc2 present in the system relative to the system without the addition of the composition.

99. Also disclosed are methods of identifying inhibitors of HIV budding comprising, a) administering a composition to a system, wherein the system supports HIV budding via a Vps4a-Bc2 interaction, b) assaying the effect of the composition on the amount of HIV budding in the system, and c) selecting a composition which causes a decrease in the amount of HIV budding present in the system because of an inhibition of the Vps4a-Bc2 interaction relative to the system without the addition of the composition.

100. Disclosed are methods of inhibiting HIV budding comprising administering a composition, wherein the composition prevents HIV budding, wherein the composition is defined as a composition capable of being identified by administering the composition to a system, wherein the system supports HIV-budding via a Vps4a-Bc2 interaction, assaying the effect of the composition on the amount of HIV budding in the system, and selecting a composition which causes a decrease in the amount of HIV budding present in the system because of an inhibition of the Vps4a-Bc2 interaction relative to the system without the addition of the composition.

101. Also disclosed are methods of inhibiting HIV budding comprising administering a composition that reduces an interaction between Vps4a-Bc2.

102. Disclosed are methods of making a composition capable of inhibiting HIV budding comprising admixing a compound with a pharmaceutically acceptable carrier, wherein the compound is identified by administering the compound to a system, wherein the system supports HIV budding via a

Vps4a-Bc2interaction, assaying the effect of the compound on the amount of HIV budding in the system, and selecting a compound which causes a decrease in the amount of HIV budding in the system because of an inhibition of the Vps4a-Bc2interaction, relative to the system without the addition of the compound.

103. Disclosed are methods of manufacturing an inhibitor to HIV budding comprising, a) administering a composition to a system, wherein the system supports HIV budding via a Vps4a-Bc2 interaction, b) assaying the effect of the composition on the amount of HIV budding in the system, c) selecting a composition which causes a decrease in the amount of HIV budding present in the system because of an inhibition of the Vps4a-Bc2 interaction, relative to the system with the addition of the composition, and d) synthesizing the composition.

104. Also disclosed are methods comprising the step of admixing the composition with a pharmaceutical carrier.

105. Disclosed are methods of identifying an inhibitor of an interaction between Vps4a and Bc2 comprising a) administering a composition to a system, wherein the system comprises Vps4a, b) assaying the effect of the composition on a Vps4a-Bc2 interaction, and c) selecting a composition which inhibits a Vps4a-Bc2 interaction.

106. Disclosed are cells that further comprising an inhibitor of a Vps4a-Bc2 interaction or an inhibitor of Vps4a-Bc2 induced viral budding.

5. Vps4b-Bc2 interactions and effects

107. Disclosed are methods for reducing interactions between Vps4b and Bc2, comprising incubating an inhibitor of the interaction between Vps4b and Bc2. Also disclosed are methods for inhibiting retroviral budding comprising administering an inhibitor of the interaction between Vps4b and Bc2 or administration of Bc2, such as excess Bc2.

108. Disclosed are methods of treating a subject comprising administering to the subject an inhibitor of HIV budding, wherein the inhibitor reduces the interaction between Vps4b and Bc2, and wherein the subject is in need of such treatment, and wherein the inhibitor is Bc2 or analog.

109. Also disclosed are methods, wherein the inhibitor disrupts an interaction between Vps4b and the Bc2.

110. Disclosed are methods of identifying an inhibitor of an interaction between Vps4b and Bc2 comprising incubating a library of molecules with Vps4b forming a mixture, and identifying the molecules that disrupt the interaction between Vps4b and Bc2, wherein the interaction disrupted comprises an interaction between the Bc2 and an amino acid of Vps4b.

111. Also disclosed are methods, wherein the step of isolating comprises incubating the mixture with a molecule comprising Bc2 or fragment thereof.

112. Disclosed are methods of identifying an inhibitor of an interaction between Vps4b and Bc2 comprising incubating a library of molecules with Bc2 forming a mixture, and identifying the molecules that disrupt the interaction between Vps4b and Bc2, wherein the interaction disrupted comprises an interaction between the Bc2 and an amino acid of Vps4b.

113. Also disclosed are methods, wherein the step of isolating comprises incubating the mixture with a molecule comprising Vps4b or fragment thereof.

114. Also disclosed are compositions produced by any of the processes as disclosed herein, as well as compositions capable of being identified by the processes disclosed herein.

115. It is understood that the disclosed methods and compositions are active for retroviruses dependent on GAG, such as HIV and Ebola.

116. Disclosed are methods of manufacturing a composition for inhibiting the interaction between Vps4b and Bc2 comprising synthesizing the inhibitors as disclosed herein.

117. Also disclosed are methods that include mixing a pharmaceutical carrier with the inhibitor as disclosed herein, and produced by any of the disclosed methods.

118. Disclosed are methods of identifying inhibitors of Vps4b and Bc2 interaction comprising, a) administering a composition to a system, wherein the system supports Vps4b-Bc2 interaction, b) assaying the effect of the composition on the amount of Vps4b-Bc2 is in the system, and c) selecting a composition which causes a decrease in the amount of Vps4b-Bc2 present in the system relative to the system without the addition of the composition.

119. Also disclosed are methods of identifying inhibitors of HIV budding comprising, a) administering a composition to a system, wherein the system supports HIV budding via a Vps4b-Bc2 interaction, b) assaying the effect of the composition on the amount of HIV budding in the system, and c) selecting a composition which causes a decrease in the amount of HIV budding present in the system because of an inhibition of the Vps4b-Bc2 interaction relative to the system without the addition of the composition.

120. Disclosed are methods of inhibiting HIV budding comprising administering a composition, wherein the composition prevents HIV budding, wherein the composition is defined as a composition capable of being identified by administering the composition to a system, wherein the system supports HIV-budding via a Vps4b-Bc2 interaction, assaying the effect of the composition on the amount of HIV budding in the system, and selecting a composition which causes a decrease in the amount of HIV budding present in the system because of an inhibition of the Vps4b-Bc2 interaction relative to the system without the addition of the composition.

121. Also disclosed are methods of inhibiting HIV budding comprising administering a composition that reduces an interaction between Vps4b-Bc2.

122. Disclosed are methods of making a composition capable of inhibiting HIV budding comprising admixing a compound with a pharmaceutically acceptable carrier, wherein the compound is identified by administering the compound to a system, wherein the system supports HIV budding via a Vps4b-Bc2interaction, assaying the effect of the compound on the amount of HIV budding in the system, and selecting a compound which causes a decrease in the amount of HIV budding in the system because of an inhibition of the Vps4b-Bc2interaction, relative to the system without the addition of the compound.

123. Disclosed are methods of manufacturing an inhibitor to HIV budding comprising, a) administering a composition to a system, wherein the system supports HIV budding via a Vps4b-Bc2 interaction, b) assaying the effect of the composition on the amount of HIV budding in the system, c) selecting a composition which causes a decrease in the amount of HIV budding present in the system because of an inhibition of the Vps4b-Bc2 interaction, relative to the system with the addition of the composition, and d) synthesizing the composition.

124. Also disclosed are methods comprising the step of admixing the composition with a pharmaceutical carrier.

125. Disclosed are methods of identifying an inhibitor of an interaction between Vps4b and Bc2 comprising a) administering a composition to a system, wherein the system comprises Vps4b, b) assaying the effect of the composition on a Vps4b-Bc2 interaction, and c) selecting a composition which inhibits a Vps4b-Bc2 interaction.

126. Disclosed are cells that further comprising an inhibitor of a Vps4b-Bc2 interaction or an inhibitor of Vps4b-Bc2 induced viral budding.

6. TSG101-GAG interactions and effects

127. Disclosed are methods for reducing interactions between TSG101 and HIV GAG, comprising incubating an inhibitor of the interaction between TSG101 and GAG with TSG101, GAG, or TSG101 and HIV GAG.

128. Also disclosed are methods for inhibiting HIV budding comprising administering an inhibitor of the interaction between TSG101 and HIV GAG.

129. Disclosed are methods of treating a subject comprising administering to the subject an inhibitor of HIV budding.

130. Disclosed are methods wherein the inhibitor prevents expression of TSG101 mRNA, wherein the inhibitor interacts with the mRNA of TSG10, wherein the inhibitor is an interfering RNA.

131. Disclosed are methods for inhibiting viral budding in a virus comprising proteins that interact with TSG101, comprising administering an inhibitor of the Vps1 or Vps 4A or 4B pathway in a cell.

132. Disclosed are methods, wherein the virus is HIV or Ebola.

133. Disclosed are methods, wherein the inhibitor prevents normal trafficking of TSG101, wherein the inhibitor reduces the release of TSG101 from endosomes, wherein the inhibitor prevents normal Vps4A or 4B function, wherein the inhibitor interacts with Vps4A or 4B.

134. Disclosed are methods for reducing interactions between TSG101 and HIV GAG, comprising incubating an inhibitor of the interaction between TSG101 and GAG. Also disclosed are methods for inhibiting HIV budding comprising administering an inhibitor of the interaction between TSG101 and HIV GAG.

135. Disclosed are methods of treating a subject comprising administering to the subject an inhibitor of HIV budding, wherein the inhibitor reduces the interaction between TSG101 and HIV GAG, and wherein the subject is in need of such treatment.

136. Also disclosed are methods, wherein the HIV-GAG comprises a P6 region, wherein the inhibitor disrupts an interaction between TSG101 and the P6 region of HIV GAG, wherein the HIV-GAG comprises a PTAP domain, wherein the inhibitor disrupts an interaction between TSG101 and the PTAP domain, and/or wherein the inhibitor interacts with the UEV domain of TSG101.

137. Disclosed are methods of identifying an inhibitor of an interaction between TSG101 and retroviral GAG comprising incubating a library of molecules with TSG101 forming a mixture, and identifying the molecules that disrupt the interaction between TSG101 and retroviral GAG, wherein the interaction disrupted comprises an interaction between the retroviral GAG and an amino acid of TSG101.

138. Disclosed are methods of identifying an inhibitor of Vsp4 ATPase activity comprising incubating a library of molecules with ATPase forming a mixture, and identifying the molecules that inhibit the ATPase activity. Also disclosed are methods wherein the ATPase activity is determined through assaying viral budding, such as HIV viral budding, as disclosed herein.

139. Also disclosed are methods, wherein the step of isolating comprises incubating the mixture with a molecule comprising a PTAP domain.

140. Disclosed are methods of identifying an inhibitor of an interaction between TSG101 and retroviral GAG comprising incubating a library of molecules with retroviral GAG forming a mixture, and identifying the molecules that disrupt the interaction between retroviral GAG and TSG101, wherein the interaction disrupted comprises an interaction between the retroviral GAG and an amino acid of TSG101.

141. Also disclosed are the methods, wherein the step of isolating comprises incubating the mixture with a molecule comprising a TSG 101 UEV domain.

142. Also disclosed are compositions produced by any of the processes as disclosed herein, as well as compositions capable of being identified by the processes disclosed herein.

143. It is understood that the disclosed methods and compositions are active for retroviruses dependent on GAG, such as HIV and Ebola.

144. It is also understood that the disclosed methods and compositions can be further combined, with for example, Ubiquitin addition. Disclosed herein, Ub and PTAP domains bind TSG101 such that the binding of the two in the presence of each other is tight then the binding of either alone. Thus, disclosed are methods and assays where, for example, both PTAP domains and Ub are incubated with target molecules or used in systems to assays for molecules that inhibit retroviral budding. Also disclosed are tripartite systems, wherein Ub, TSG101 and GAG or fragments thereof are present or utilized in the disclosed method in any of the possible permutations.

145. Disclosed are methods of manufacturing a composition for inhibiting the interaction between TSG101 and GAG comprising synthesizing the inhibitors as disclosed herein.

146. Also disclosed are methods that include mixing a pharmaceutical carrier with the inhibitor as disclosed herein, and produced by any of the disclosed methods.

147. Disclosed are methods of identifying inhibitors of TSG101-GAG interaction comprising, a) administering a composition to a system, wherein the system supports TSG101-GAG interaction, b) assaying the effect of the composition on the amount of TSG101-GAG is in the system, and c) selecting a composition which causes a decrease in the amount of TSG101-GAG present in the system relative to the system without the addition of the composition.

148. Also disclosed are methods of identifying inhibitors of HIV budding comprising, a) administering a composition to a system, wherein the system supports HIV budding via a TSG101-GAG interaction, b) assaying the effect of the composition on the amount of HIV budding in the system, and c) selecting a composition which causes a decrease in the amount of HIV budding present in the system because of an inhibition of the TSG101-GAG interaction relative to the system without the addition of the composition.

149. Disclosed are methods of inhibiting HIV budding comprising administering a composition, wherein the composition prevents HIV budding, wherein the composition is defined as a composition capable of being identified by administering the composition to a system, wherein the system supports HIV-budding via a TSG101-GAG interaction, assaying the effect of the composition on the amount of HIV budding in the system, and selecting a composition which causes a decrease in the amount of HIV budding present in the system because of an inhibition of the TSG101-GAG interaction relative to the system without the addition of the composition.

150. Also disclosed are methods of inhibiting HIV budding comprising administering a composition that reduces an interaction between TSG101 and GAG.

151. Disclosed are methods of making a composition capable of inhibiting HIV budding comprising admixing a compound with a pharmaceutically acceptable carrier, wherein the compound is identified by administering the compound to a system, wherein the system supports HIV budding via a TSG101-GAG interaction, assaying the effect of the compound on the amount of HIV budding in the system, and selecting a compound which causes a decrease in the amount of HIV budding in the system because of an inhibition of the TSG101-GAG interaction, relative to the system without the addition of the compound.

152. Disclosed are methods of manufacturing an inhibitor to HIV budding comprising, a) administering a composition to a system, wherein the system supports HIV budding via a TSG101-GAG interaction, b) assaying the effect of the composition on the amount of HIV budding in the system, c) selecting a composition which causes a decrease in the amount of HIV budding present in the system because of an inhibition of the TSG101-GAG interaction, relative to the system with the addition of the composition, and d) synthesizing the composition.

153. Also disclosed are methods comprising the step of admixing the composition with a pharmaceutical carrier.

154. Disclosed are methods of identifying an inhibitor of an interaction between TSG101 and GAG comprising a) administering a composition to a system, wherein the system comprises TSG101, b) assaying the effect of the composition on a TSG101-GAG interaction, and c) selecting a composition which inhibits a TSG101-GAG interaction.

155. Disclosed are methods, wherein the GAG comprises a P6 region, wherein the inhibitor disrupts an interaction between TSG101 and the P6 region of GAG, and/or wherein the GAG comprises a PTAP domain, wherein the inhibitor disrupts an interaction between TSG101 and the PTAP domain.

156. Disclosed are cells comprising, a) a regulatable nucleic acid comprising sequence encoding a TSG101 gene and b) a nucleic acid comprising sequence encoding a GAG gene as well as cells comprising a regulatable nucleic acid comprising sequence encoding a TSG101 gene and sequence comprising a GAG gene.

157. Disclosed are cells that further comprising an inhibitor of a TSG101-GAG interaction.

7. TSG101-Vps28 interactions and effects

158. Disclosed are methods for reducing interactions between TSG101 and Vps28, comprising incubating an inhibitor of the interaction between TSG101 and Vps28. Also disclosed are methods for inhibiting retroviral budding comprising administering an inhibitor of the interaction between TSG101 and Vps28.

159. Disclosed are methods of treating a subject comprising administering to the subject an inhibitor of HIV budding, wherein the inhibitor reduces the interaction between TSG101 and Vps28, and wherein the subject is in need of such treatment.

160. Also disclosed are methods, wherein the inhibitor disrupts an interaction between TSG101 and the Vps28, and/or wherein the inhibitor interacts with the UEV domain of TSG101.

161. Disclosed are methods of identifying an inhibitor of an interaction between TSG101 and Vps28 comprising incubating a library of molecules with TSG101 forming a mixture, and identifying the molecules that disrupt the interaction between TSG101 and Vps28, wherein the interaction disrupted comprises an interaction between the Vps28 and an amino acid of TSG101.

162. Also disclosed are methods, wherein the step of isolating comprises incubating the mixture with a molecule comprising Vps28.

163. Disclosed are methods of identifying an inhibitor of an interaction between TSG101 and Vps28 comprising incubating a library of molecules with Vps28 forming a mixture, and identifying the molecules that disrupt the interaction between Vps28 and TSG101, wherein the interaction disrupted comprises an interaction between the Vps28 and an amino acid of TSG101.

164. Also disclosed are the methods, wherein the step of isolating comprises incubating the mixture with a molecule comprising a TSG 101 UEV domain.

165. Also disclosed are compositions produced by any of the processes as disclosed herein, as well as compositions capable of being identified by the processes disclosed herein.

166. It is understood that the disclosed methods and compositions are active for retroviruses dependent on GAG, such as HIV and Ebola.

167. It is also understood that the disclosed methods and compositions can be further combined, with for example, PTAP domain addition. Disclosed herein, Vps28 and PTAP domains bind TSG101 such that the binding of the two in the presence of each other is tight then the binding of either alone. Thus, disclosed are methods and assays where, for example, both PTAP domains and Vps28 are incubated with target molecules or used in systems to assays for molecules that inhibit retroviral budding.

168. Disclosed are methods of manufacturing a composition for inhibiting the interaction between TSG101 and Vps28 comprising synthesizing the inhibitors as disclosed herein.

169. Also disclosed are methods that include mixing a pharmaceutical carrier with the inhibitor as disclosed herein, and produced by any of the disclosed methods.

170. Disclosed are methods of identifying inhibitors of TSG101-Vps28 interaction comprising, a) administering a composition to a system, wherein the system supports TSG101-Vps28 interaction, b) assaying the effect of the composition on the amount of TSG101-Vps28 is in the system, and c) selecting a composition which causes a decrease in the amount of TSG101-Vps28 present in the system relative to the system without the addition of the composition.

171. Also disclosed are methods of identifying inhibitors of HIV budding comprising, a) administering a composition to a system, wherein the system supports HIV budding via a TSG101-Vps28 interaction, b) assaying the effect of the composition on the amount of HIV budding in the system, and c) selecting a composition which causes a decrease in the amount of HIV budding present in the system because of an inhibition of the TSG101-Vps28 interaction relative to the system without the addition of the composition.

172. Disclosed are methods of inhibiting HIV budding comprising administering a composition, wherein the composition prevents HIV budding, wherein the composition is defined as a composition capable of being identified by administering the composition to a system, wherein the system supports HIV-budding via a TSG101-Vps28 interaction, assaying the effect of the composition on the amount of HIV budding in the system, and selecting a composition which causes a decrease in the amount of HIV budding present in the system because of an inhibition of the TSG101-Vps28 interaction relative to the system without the addition of the composition.

173. Also disclosed are methods of inhibiting HIV budding comprising administering a composition that reduces an interaction between TSG101 and Vps28.

174. Disclosed are methods of making a composition capable of inhibiting HIV budding comprising admixing a compound with a pharmaceutically acceptable carrier, wherein the compound is identified by administering the compound to a system, wherein the system supports HIV budding via a TSG101-Vps28 interaction, assaying the effect of the compound on the amount of HIV budding in the system, and selecting a compound which causes a decrease in the amount of HIV budding in the system

because of an inhibition of the TSG101-Vps28 interaction, relative to the system without the addition of the compound.

175. Disclosed are methods of manufacturing an inhibitor to HIV budding comprising, a) administering a composition to a system, wherein the system supports HIV budding via a TSG101-Vps28 interaction, b) assaying the effect of the composition on the amount of HIV budding in the system, c) selecting a composition which causes a decrease in the amount of HIV budding present in the system because of an inhibition of the TSG101-Vps28 interaction, relative to the system with the addition of the composition, and d) synthesizing the composition.

176. Also disclosed are methods comprising the step of admixing the composition with a pharmaceutical carrier.

177. Disclosed are methods of identifying an inhibitor of an interaction between TSG101 and Vps28 comprising a) administering a composition to a system, wherein the system comprises TSG101, b) assaying the effect of the composition on a TSG101-Vps28 interaction, and c) selecting a composition which inhibits a TSG101-Vps28 interaction.

178. Disclosed are cells that further comprising an inhibitor of a TSG101-Vps28 interaction.

179. TSG101-Vps37 interactions and effects

180. Disclosed are methods for reducing interactions between TSG101 and Vps37, comprising incubating an inhibitor of the interaction between TSG101 and Vps37. Also disclosed are methods for inhibiting retroviral budding comprising administering an inhibitor of the interaction between TSG101 and Vps37.

181. Disclosed are methods of treating a subject comprising administering to the subject an inhibitor of HIV budding, wherein the inhibitor reduces the interaction between TSG101 and Vps37, and wherein the subject is in need of such treatment.

182. Also disclosed are methods, wherein the inhibitor disrupts an interaction between TSG101 and the Vps37, and/or wherein the inhibitor interacts with the UEV domain of TSG101.

183. Disclosed are methods of identifying an inhibitor of an interaction between TSG101 and Vps37 comprising incubating a library of molecules with TSG101 forming a mixture, and identifying the molecules that disrupt the interaction between TSG101 and Vps37, wherein the interaction disrupted comprises an interaction between the Vps37 and an amino acid of TSG101.

184. Also disclosed are methods, wherein the step of isolating comprises incubating the mixture with a molecule comprising Vps37.

185. Disclosed are methods of identifying an inhibitor of an interaction between TSG101 and Vps37 comprising incubating a library of molecules with Vps37 forming a mixture, and identifying the molecules that disrupt the interaction between Vps37 and TSG101, wherein the interaction disrupted comprises an interaction between the Vps37 and an amino acid of TSG101.

186. Also disclosed are the methods, wherein the step of isolating comprises incubating the mixture with a molecule comprising a TSG 101 UEV domain.

187. Also disclosed are compositions produced by any of the processes as disclosed herein, as well as compositions capable of being identified by the processes disclosed herein.

188. It is understood that the disclosed methods and compositions are active for retroviruses dependent on GAG, such as HIV and Ebola.

189. It is also understood that the disclosed methods and compositions can be further combined, with for example, PTAP domain addition. Disclosed herein, Vps37 and PTAP domains bind TSG101 such that the binding of the two in the presence of each other is tight then the binding of either alone. Thus, disclosed are methods and assays where, for example, both PTAP domains and Vps37 are incubated with target molecules or used in systems to assays for molecules that inhibit retroviral budding.

190. Disclosed are methods of manufacturing a composition for inhibiting the interaction between TSG101 and Vps37 comprising synthesizing the inhibitors as disclosed herein.

191. Also disclosed are methods that include mixing a pharmaceutical carrier with the inhibitor as disclosed herein, and produced by any of the disclosed methods.

192. Disclosed are methods of identifying inhibitors of TSG101-Vps37 interaction comprising, a) administering a composition to a system, wherein the system supports TSG101-Vps37 interaction, b) assaying the effect of the composition on the amount of TSG101-Vps37 is in the system, and c) selecting a composition which causes a decrease in the amount of TSG101-Vps37 present in the system relative to the system without the addition of the composition.

193. Also disclosed are methods of identifying inhibitors of HIV budding comprising, a) administering a composition to a system, wherein the system supports HIV budding via a TSG101-Vps37 interaction, b) assaying the effect of the composition on the amount of HIV budding in the system, and c) selecting a composition which causes a decrease in the amount of HIV budding present in the system because of an inhibition of the TSG101-Vps37 interaction relative to the system without the addition of the composition.

194. Disclosed are methods of inhibiting HIV budding comprising administering a composition, wherein the composition prevents HIV budding, wherein the composition is defined as a composition capable of being identified by administering the composition to a system, wherein the system supports HIV-budding via a TSG101-Vps37 interaction, assaying the effect of the composition on the amount of HIV budding in the system, and selecting a composition which causes a decrease in the amount of HIV budding present in the system because of an inhibition of the TSG101-Vps37 interaction relative to the system without the addition of the composition.

195. Also disclosed are methods of inhibiting HIV budding comprising administering a composition that reduces an interaction between TSG101 and Vps37.

196. Disclosed are methods of making a composition capable of inhibiting HIV budding comprising admixing a compound with a pharmaceutically acceptable carrier, wherein the compound is identified by administering the compound to a system, wherein the system supports HIV budding via a TSG101-Vps37 interaction, assaying the effect of the compound on the amount of HIV budding in the

system, and selecting a compound which causes a decrease in the amount of HIV budding in the system because of an inhibition of the TSG101-Vps37 interaction, relative to the system without the addition of the compound.

197. Disclosed are methods of manufacturing an inhibitor to HIV budding comprising, a) administering a composition to a system, wherein the system supports HIV budding via a TSG101-Vps37 interaction, b) assaying the effect of the composition on the amount of HIV budding in the system, c) selecting a composition which causes a decrease in the amount of HIV budding present in the system because of an inhibition of the TSG101-Vps37 interaction, relative to the system with the addition of the composition, and d) synthesizing the composition.

198. Also disclosed are methods comprising the step of admixing the composition with a pharmaceutical carrier.

199. Disclosed are methods of identifying an inhibitor of an interaction between TSG101 and Vps37 comprising a) administering a composition to a system, wherein the system comprises TSG101, b) assaying the effect of the composition on a TSG101-Vps37 interaction, and c) selecting a composition which inhibits a TSG101-Vps37 interaction.

200. Disclosed are cells that further comprise an inhibitor of a TSG101-Vps37 interaction.

8. Escrt 1 complex interactions and effects

201. Disclosed are methods for reducing interactions within the Escrt 1 complex, comprising incubating an inhibitor of the interaction with Escrt 1 complex. Also disclosed are methods for inhibiting retroviral budding comprising administering an inhibitor of the interaction within the Escrt 1 complex.

202. Disclosed are methods of treating a subject comprising administering to the subject an inhibitor of HIV budding, wherein the inhibitor reduces the interaction within the Escrt 1 complex, and wherein the subject is in need of such treatment.

203. Also disclosed are methods, wherein the inhibitor disrupts an interaction within the Escrt 1 complex, and/or wherein the inhibitor interacts with the UEV domain of TSG101, and/or S-box domain orf TSG101, and/or the PTAP domain of TSG101.

204. Disclosed are methods of identifying an inhibitor of an interaction within the Escrt 1 complex comprising incubating a library of molecules with Escrt 1 complex forming a mixture, and identifying the molecules that disrupt the interaction within the Escrt 1 complex, wherein the interaction disrupted comprises an interaction between an amino acid of a protein within the Escrt 1 complex.

205. Also disclosed are methods, wherein the step of isolating comprises incubating the mixture with a molecule comprising a molecule comprising a protein within the Escrt 1 complex or the Escrt 1 complex itself.

206. Disclosed are methods of identifying an inhibitor of an interaction within the Escrt 1 complex comprising incubating a library of molecules with the Escrt 1 complex forming a mixture, and identifying the molecules that disrupt the interaction within the Escrt 1 complex, wherein the interaction disrupted comprises an interaction with an amino acid of a protein within the Escrt 1 complex.

207. Also disclosed are the methods, wherein the step of isolating comprises incubating the mixture with a molecule comprising a TSG 101 UEV domain, and/or a TSG101-PTAP domain, and/or a TSG101-S-Box domain.

208. Also disclosed are compositions produced by any of the processes as disclosed herein, as well as compositions capable of being identified by the processes disclosed herein.

209. It is understood that the disclosed methods and compositions are active for retroviruses dependent on GAG, such as HIV and Ebola.

210. It is also understood that the disclosed methods and compositions can be further combined, with ay of the disclosed methods and compositions herein.

211. Disclosed are methods of manufacturing a composition for inhibiting the interaction within the Escrt 1 complex comprising synthesizing the inhibitors as disclosed herein.

212. Also disclosed are methods that include mixing a pharmaceutical carrier with the inhibitor as disclosed herein, and produced by any of the disclosed methods.

213. Disclosed are methods of identifying inhibitors of an interaction within a Escrt 1 complex comprising, a) administering a composition to a system, wherein the system supports Escrt 1 complex formation, b) assaying the effect of the composition on the amount of Escrt 1 complex formed in the system, and c) selecting a composition which causes a decrease in the amount of Escrt 1 complex present in the system relative to the system without the addition of the composition.

214. Also disclosed are methods of identifying inhibitors of HIV budding comprising, a) administering a composition to a system, wherein the system supports HIV budding via Escrt 1 complex formation, b) assaying the effect of the composition on the amount of HIV budding in the system, and c) selecting a composition which causes a decrease in the amount of HIV budding present in the system because of an inhibition of the Escrt 1 complex relative to the system without the addition of the composition.

215. Disclosed are methods of inhibiting HIV budding comprising administering a composition, wherein the composition prevents HIV budding, wherein the composition is defined as a composition capable of being identified by administering the composition to a system, wherein the system supports HIV-budding via a TSG101-Vps37 interaction, assaying the effect of the composition on the amount of HIV budding in the system, and selecting a composition which causes a decrease in the amount of HIV budding present in the system because of an inhibition of the TSG101-Vps37 interaction relative to the system without the addition of the composition.

216. Also disclosed are methods of inhibiting HIV budding comprising administering a composition that reduces an interaction within the Escrt 1 complex.

217. Disclosed are methods of making a composition capable of inhibiting HIV budding comprising admixing a compound with a pharmaceutically acceptable carrier, wherein the compound is identified by administering the compound to a system, wherein the system supports HIV budding via a TSG101-Vps37 interaction, assaying the effect of the compound on the amount of HIV budding in the

system, and selecting a compound which causes a decrease in the amount of HIV budding in the system because of an inhibition of the TSG101-Vps37 interaction, relative to the system without the addition of the compound.

218. Disclosed are methods of manufacturing an inhibitor to HIV budding comprising, a) administering a composition to a system, wherein the system supports HIV budding via an Escrt 1 complex formation, b) assaying the effect of the composition on the amount of HIV budding in the system, c) selecting a composition which causes a decrease in the amount of HIV budding present in the system because of an inhibition of the Escrt 1 complex formation, relative to the system with the addition of the composition, and d) synthesizing the composition.

219. Also disclosed are methods comprising the step of admixing the composition with a pharmaceutical carrier.

220. Disclosed are methods of identifying an inhibitor of an interaction within the Escrt 1 complex comprising a) administering a composition to a system, wherein the system comprises an Escrt 1 complex, b) assaying the effect of the composition on the Escrt 1 complex, and c) selecting a composition which inhibits a Escrt 1 complex formation.

221. Disclosed are cells that further comprise an inhibitor of the Escrt complex formation.

9. Downstream Escrt complex interactions and effects

222. Disclosed are methods for reducing interactions within the Downstream Escrt complex, comprising incubating an inhibitor of the interaction with the Downstream Escrt complex. Also disclosed are methods for inhibiting retroviral budding comprising administering an inhibitor of the interaction within the Downstream Escrt complex.

223. Disclosed are methods of treating a subject comprising administering to the subject an inhibitor of HIV budding, wherein the inhibitor reduces the interaction within the Downstream Escrt complex, and wherein the subject is in need of such treatment.

224. Also disclosed are methods, wherein the inhibitor disrupts an interaction within the Downstream Escrt complex, and/or wherein the inhibitor interacts with the UEV domain of TSG101, and.or S-box domain orf TSG101, and/or the PTAP domain of TSG101.

225. Disclosed are methods of identifying an inhibitor of an interaction within the Downstream Escrt complex comprising incubating a library of molecules with the Downstream Escrt complex forming a mixture, and identifying the molecules that disrupt the interaction within the Downstream Escrt complex, wherein the interaction disrupted comprises an interaction between an amino acid of a protein within the Downstream Escrt complex.

226. Also disclosed are methods, wherein the step of isolating comprises incubating the mixture with a molecule comprising a molecule comprising a protein within the Downstream Escrt complex or the Downstream Escrt complex itself.

227. Disclosed are methods of identifying an inhibitor of an interaction within the Downstream Escrt complex comprising incubating a library of molecules with the Downstream Escrt complex

forming a mixture, and identifying the molecules that disrupt the interaction within the Downstream Escrt complex, wherein the interaction disrupted comprises an interaction with an amino acid of a protein within the Downstream Escrt complex.

228. Also disclosed are the methods, wherein the step of isolating comprises incubating the mixture with a molecule comprising a TSG 101 UEV domain, and/or a TSG101-PTAP domain, and/or a TSG101-S-Box domain.

229. Also disclosed are compositions produced by any of the processes as disclosed herein, as well as compositions capable of being identified by the processes disclosed herein.

230. It is understood that the disclosed methods and compositions are active for retroviruses dependent on GAG, such as HIV and Ebola.

231. It is also understood that the disclosed methods and compositions can be further combined, with ay of the disclosed methods and compositions herein.

232. Disclosed are methods of manufacturing a composition for inhibiting the interaction within the Downstream Escrt complex comprising synthesizing the inhibitors as disclosed herein.

233. Also disclosed are methods that include mixing a pharmaceutical carrier with the inhibitor as disclosed herein, and produced by any of the disclosed methods.

234. Disclosed are methods of identifying inhibitors of an interaction within a Downstream Escrt complex comprising, a) administering a composition to a system, wherein the system supports Downstream Escrt complex formation, b) assaying the effect of the composition on the amount of Downstream Escrt complex formed in the system, and c) selecting a composition which causes a decrease in the amount of Downstream Escrt complex present in the system relative to the system without the addition of the composition.

235. Also disclosed are methods of identifying inhibitors of HIV budding comprising, a) administering a composition to a system, wherein the system supports HIV budding via Downstream Escrt complex formation, b) assaying the effect of the composition on the amount of HIV budding in the system, and c) selecting a composition which causes a decrease in the amount of HIV budding present in the system because of an inhibition of the Downstream Escrt complex relative to the system without the addition of the composition.

236. Disclosed are methods of inhibiting HIV budding comprising administering a composition, wherein the composition prevents HIV budding, wherein the composition is defined as a composition capable of being identified by administering the composition to a system, wherein the system supports HIV-budding via a TSG101-Vps37 interaction, assaying the effect of the composition on the amount of HIV budding in the system, and selecting a composition which causes a decrease in the amount of HIV budding present in the system because of an inhibition of the TSG101-Vps37 interaction relative to the system without the addition of the composition.

237. Also disclosed are methods of inhibiting HIV budding comprising administering a composition that reduces an interaction within the Downstream Escrt complex.

238. Disclosed are methods of making a composition capable of inhibiting HIV budding comprising admixing a compound with a pharmaceutically acceptable carrier, wherein the compound is identified by administering the compound to a system, wherein the system supports HIV budding via a TSG101-Vps37 interaction, assaying the effect of the compound on the amount of HIV budding in the system, and selecting a compound which causes a decrease in the amount of HIV budding in the system because of an inhibition of the TSG101-Vps37 interaction, relative to the system without the addition of the compound.

239. Disclosed are methods of manufacturing an inhibitor to HIV budding comprising, a) administering a composition to a system, wherein the system supports HIV budding via an Downstream Escrt complex formation, b) assaying the effect of the composition on the amount of HIV budding in the system, c) selecting a composition which causes a decrease in the amount of HIV budding present in the system because of an inhibition of the Downstream Escrt complex formation, relative to the system with the addition of the composition, and d) synthesizing the composition.

240. Also disclosed are methods comprising the step of admixing the composition with a pharmaceutical carrier.

241. Disclosed are methods of identifying an inhibitor of an interaction within the Downstream Escrt complex comprising a) administering a composition to a system, wherein the system comprises an Downstream Escrt complex, b) assaying the effect of the composition on the Downstream Escrt complex, and c) selecting a composition which inhibits a Downstream Escrt complex formation.

242. Disclosed are cells that further comprise an inhibitor of a Escrt complex formation.

10. General Interactions

243. Disclosed herein are a variety of molecules and a variety of compositions and a variety of methods. It is understood that the general principles disclosed by these methods can be applied to any molecule or set of molecules or composition or set of compositions disclosed herein. For example, the disclosed methods could be applied to TSG101-Vps4a or Vps4b interactions, or they could be applied to any interaction with Ub, such as a Ub-Vps4a interaction, a Ub-Vps4b interaction, a Ub-GAG interaction, as well as Vps28-Vps4a or Vps4b interactions, or any interaction between a protein or molecule in the Escrt 1 complex and any molecule in the downstream Escrt complex. It is also understood that a variety of functions exist for these molecules and compositions disclosed herein, such as binding, promotion of viral budding, promotion of GAG maturation, or ATPase function. There are number of assays disclosed herein related to these functions, including binding assays, protein production assays, and viral infectivity assays. Also it is understood that methods and compositions regarding, for example, Bc2-Ub, Bc2-GAG, Vsp4a-TSG101, Ub-GAG are disclosed.

C. Compositions

1. TSG101

244. TSG101 comprises a number of domains. TSG101 can have a PTAP binding domain, which functions to interact with the PTAP region of a retroviral GAG protein. It is understood that the TSG101 PTAP binding domain can bind any protein containing the PTAP region. This region is a part of

retroviral GAG proteins as well as proteins in enveloped viruses TSG101 also can have ubiquitin binding domain, which can interact with ubiquitin.

245. The direct interaction between Tsg101 and the PTAP late domain motifs on the viral structural proteins is mediated by the N-terminal ubiquitin E2 variant (UEV) domain of Tsg101 (VerPlank et al., 2001; Garrus et al., 2001; Martin-Serrano et al., 2001), whereas the remaining C-terminal two-thirds of Tsg101 interacts with other proteins (Fig. 1*a*). Like other UEV domains, Tsg101 UEV shows significant sequence similarity to E2 ubiquitin ligases but is unable to catalyze ubiquitin transfer, as it lacks the active site cysteine that forms the transient thioester bond with the C-terminus of ubiquitin (Ub) (Koonin and Abagyan, 1997; Ponting et al., 1997) (Fig. 1*b*). Nevertheless, at least some UEVs have retained the ability to bind Ub, and appear to act either as co-factors in ubiquitination reactions, or as ubiquitin sensors (Hofmann and Pickart, 1999; VerPlank et al., 2001; Katzmann et al., 2001; VanDemark et al., 2001; Garrus et al., 2001). UEV domains also frequently contain other protein recognition motifs, and may generally serve to couple protein and Ub binding functions to facilitate the formation of multiprotein complexes.

246. Tsg101 normally functions in the cellular vacuolar protein sorting (Vps) pathway, which coordinates the sorting of membrane-associated proteins through a series of endosomal compartments for eventual delivery to the lysosome (vacuole in yeast) (Dupre et al., 2001; Hicke, 2001; Piper and Luzio, 2001). A key decision-making step in this pathway occurs when membrane patches containing proteins destined for destruction bud as small vesicles into the lumen of the late endosome, creating an organelle called the "multivesicular body" (MVB). Subsequent fusion of MVB with the lysosome delivers these vesicles and all associated cargo to the lumen of the lysosome, where they are degraded by proteolysis. In contrast, proteins that remain in the limiting membrane of MVB are delivered to the limiting membrane of lysosomes and therefore escape degradation.

247. Work, particularly in the yeast system, has revealed that Tgs101 (yeast Vps23p) performs an important role in selecting which proteins enter the MVB lumen and which remain on the limiting membrane. In yeast, proteins destined for the lumen of the vacuole are covalently modified with monoubiquitin, and bound by a 350-kDa protein complex called ESCRT-I, which is composed of Vps23p, Vps28p and Vps37p (Katzmann et al., 2001). ESCRT-I appears to function as the receptor and/or sorting complex that selects ubiquitinated proteins for incorporation into MVB vesicles. This sorting event appears absolutely dependent on the Ub-binding activity of Vps23p, as adding an ubiquitination signal targets proteins that would otherwise remain on the limiting membrane into the lumen. Conversely, a point mutation in Vps23p that abolishes Ub binding also blocks sorting of ubiquitinated cargo proteins (Katzmann et al., 2001). Although the human system is less well characterized, Tsg101 also binds Ub in vitro (Garrus et al., 2001), forms a soluble 350-kDa complex that includes Vps28 (Babst et al., 2000; Bishop and Woodman, 2001), and is required for the delivery of cathepsin D and endocytosed receptors to the lysosome (Babst et al., 2000). Thus, it appears that

ESCRT-I complex function is conserved from yeast to humans, and that recognition of ubiquitinated proteins is an important element in this function.

248. Tsg101 was initially discovered in a screen for potential tumor suppressors (Li and Cohen, 1996), and the protein appears to perform multiple functions, including down-regulating p53 via the MDM2/p53 pathway (Li et al., 2001; Ruland et al., 2001). Tsg101 also plays a central role in vacuolar protein sorting. The Vps pathway sorts membrane-bound proteins for eventual degradation in the lysosome (vacuole in yeast) ((Lemmon and Traub, 2000) and references therein). Two alternative entrées into the Vps pathway are via vesicular trafficking from the Golgi (e.g., in destroying misfolded membrane proteins) or via endocytosis from the plasma membrane (e.g., in downregulating surface receptors). Vesicles carrying proteins from either source can enter the Vps pathway by fusing with endosomes. As these endosomes mature, their cargos are sorted for lysosomal degradation via the formation of structures called multivesicular bodies (MVB). MVB are created when surface patches on late endosomes bud into the lumen, forming small (~50-100 nm) vesicles. A maturing MVB can contain tens or even hundreds of these vesicles. The MVB then fuses with the lysosome, releasing the vesicles for degradation in this hydrolytic organelle.

249. Ub also plays an important role in the budding of retroviruses and other enveloped viruses (Harty et al., 2000; Patnaik et al., 2000; Schubert et al., 2000; Strack et al., 2000; Vogt, 2000; Harty et al., 2001; Kikonyogo et al., 2001). Retroviruses contain high levels of ubiquitin, and 2-5% of Gag proteins in the virion are monoubiquitinated (Putterman et al., 1990; Ott et al., 1998; Ott et al., 2000). Moreover, treatment of infected cells with proteasome inhibitors, which decreases the intracellular concentration of free Ub, also inhibits virus release at a late stage (Schubert et al., 2000; Harty et al., 2001). In some cases this effect can be partially rescued either by overexpressing free Ub or by fusing Ub to the C-terminal end of the viral Gag protein (Patnaik et al., 2000).

2. HIV budding

250. Like other enveloped viruses, HIV-1 uses cellular machinery to bud from infected cells. Disclosed herein Tsg101 protein functions in vacuolar protein sorting (Vps) and is required for HIV-1 budding. The UEV domain of Tsg101 binds to an essential tetrapeptide (PTAP) motif within the p6 domain of the structural Gag protein and also to ubiquitin. Disclosed herein, depletion of cellular Tsg101 by a small interfering RNA arrests HIV-1 budding at a late stage, and budding is rescued by reintroduction of Tsg101. Also disclosed are dominant negative mutant Vps4 proteins, as well as other molecules, that inhibit vacuolar protein sorting also arrest HIV-1 and MLV budding. These observations are consistent with retroviruses budding by appropriating cellular machinery normally used in the Vps pathway to form multivesicular bodies.

251. HIV-1 assembly is driven by the viral Gag protein, which is actively trafficked to the plasma membrane where it associates into enveloped, spherical particles that bud from the cell (reviewed in (Freed, 1998)). During viral assembly, Gag is processed by the viral protease at a series of sites to produce four new structural proteins that perform essential functions in the mature, infectious virion

(denoted MA, CA, NC, and p6; Fig. 1A). Proteolytic processing is not required for particle production, however, as HIV-1 Gag can assemble and bud in the absence of any other viral proteins.

252. The budding of an enveloped virus can be viewed as a fission event in which the continuous cell membrane is broken and resealed to create discrete viral and cellular membranes. Like other enveloped viruses, HIV-1 does not encode its own membrane fission machinery and presumably therefore must recruit and reprogram cellular proteins to assist in the budding process. A potential docking site for such cellular factor(s) has been mapped to a conserved P(T/S)AP motif located in the p6 domain of HIV-1 Gag. Point mutations within this "PTAP" motif arrest viral release at a very late stage (Gottlinger et al., 1991; Huang et al., 1995), and it has therefore been termed a "late domain" (Wills and Craven, 1991). Virus assembly appears to initiate normally in the late domain mutants, but the continuous membrane that connects the budding particles to the cell (or to other budding particles) is not severed, resulting in abnormal or severely attenuated virus release.

253. The recruitment of cellular machinery to facilitate virus budding seems to be a general phenomenon, and distinct late domains have been identified in the structural proteins of many enveloped viruses (Vogt, 2000). Two well characterized late domains are the "PY" motif (consensus sequence: PPXY; X= any amino acid) found in membrane-associated proteins from filo-, orbi-, rhabdo- and oncoretroviruses (Craven et al., 1999; Harty et al., 2000; Harty et al., 1999; Jayakar et al., 2000), and the "YL" motif (YXXL) found in the Gag protein of equine infectious anemia virus (EIAV) (Puffer et al., 1997; Puffer et al., 1998). The various late domains can still function when moved to different positions within retroviral Gag proteins, supporting the idea that they are docking sites for cellular factors rather than structural elements (Parent et al., 1995; Yuan et al., 2000). Moreover, the different late domains can function interchangeably and multiple late domains are often found in close proximity within viral coat proteins, suggesting that they may act synergistically (Parent et al., 1995; Strack et al., 2000; Yuan et al., 2000).

254. Ubiquitin (Ub) also plays an essential, albeit poorly understood, role in retroviral budding (Patnaik et al., 2000; Schubert et al., 2000b; Strack et al., 2000; Vogt, 2000). For example, retrovirus budding can be blocked at a late stage by depleting cellular pools of free ubiquitin with proteasome inhibitors. It is not yet certain, however, whether the functionally relevant substrate for ubiquitination is Gag itself or a cellular factor. HIV-1 p6 and other retroviral Gag proteins are monoubiquitinated at low levels (Ott et al., 2000) and there is a general correlation between Gag ubiquitination and virus release (Schubert et al., 2000b; Strack et al., 2000). Moreover, the block imposed by proteasome inhibitors can be at least partially alleviated by covalently fusing Ub to the C-terminal end of Rous sarcoma virus (RSV) Gag (Patnaik et al., 2000). However, virus release and replication are not affected by mutation of the two HIV-1 p6 lysine residues (Lys-27 and Lys-33) that are the major sites for Gag ubiquitination (Ott et al., 2000). The functional role of Gag ubiquitination therefore remains uncertain. In principle, Gag ubiquitination could facilitate budding either by targeting defective Gag molecules for proteolytic

degradation and thereby preventing them from interfering with viral budding (Schubert et al., 2000a), or by creating docking sites for cellular factors that actively participate in viral budding.

3. Vps4

255. The yeast Vps4p is now known to have two 48/49kD, 80% identical orthologs in humans, Vps4a and Skd1 (Vps4b). Vps4a and Vps4b belong to a family of ATPases with more than 200 members with a conserved 200 amino acid catalytic domain called AAA (ATPases associated with various cellular activities). These domains are involved in a variety of functions, include membrane trafficking, proteasome function, the MT motor dynein, and NSF which refolds and resets SNARE complexes after membrane fusion. Most AAA ATPases form hexameric or other oligomeric rings; and Vps4 is thought to catalyze the release of membrane-bound proteins from the MVB. Vps4a andVps4b bind CHMP1 (CHarged MVB Proteins) and BC2/CHMP2 proteins. (See Babst and Emr, EMBO J. 1997, 16(8):1820-31., EMBO J. 1998, 17(11):2982-93., 1998; Bishop & Woodman, Mol Biol Cell. 2000; 11(1):227-39., J Biol Chem. 2001 Apr 13;276(15):11735-42.; and Howard et al., J. Cell Sci 2001;114(Pt 13):2395-404., all of which are herein incorporated by reference for material at least related to Vps4, Bc2, and TSG101, and other protein complexes)

256. Highly conserved lysine to glutamine and glutamate to glutamine mutations block exit of proteins from the vacuole or MVB in yeast, mouse and human cells, and thus block protein sorting in a dominant negative way. This blockage causes the MVB(or vacuole in yeast) to keep accumulating vesicles inside, which causes the organelles to swell and grow. The mutants affect ATP-binding and are K173Q, K180Q; and ATP-hydrolysis E228Q, E235Q.

4. Molecules that can interact with TSG101 and/or PTAP and/or Ub, and/or Vps4A, and/or Vps4B or Vps28 or Vps37 or Bc2 or Ubiquitin or GAG.

a) Functional Nucleic Acids

257. Functional nucleic acids are nucleic acid molecules that have a specific function, such as binding a target molecule or catalyzing a specific reaction. Functional nucleic acid molecules can be divided into the following categories, which are not meant to be limiting. For example, functional nucleic acids include antisense molecules, aptamers, ribozymes, triplex forming molecules, and external guide sequences. The functional nucleic acid molecules can act as affectors, inhibitors, modulators, and stimulators of a specific activity possessed by a target molecule, or the functional nucleic acid molecules can possess a de novo activity independent of any other molecules.

258. Functional nucleic acid molecules can interact with any macromolecule, such as DNA, RNA, polypeptides, or carbohydrate chains. Thus, functional nucleic acids can interact with the mRNA of any of the proteins and compositions disclosed herein, such as, TSG101 or Vps4A or Vps4B or fragment thereof, or the genomic DNA of any of the proteins and compositions disclosed herein, such as, TSG101 or Vps4A or Vps4B or fragment thereof or they can interact with any of the proteins and compositions disclosed herein, such as, the polypeptide TSG101 or Vps4A or Vps4B or fragment thereof, such as the PTAP binding domain or the UB binding domain. Often functional nucleic acids are designed to interact with other nucleic acids based on sequence homology between the target molecule

and the functional nucleic acid molecule. In other situations, the specific recognition between the functional nucleic acid molecule and the target molecule is not based on sequence homology between the functional nucleic acid molecule and the target molecule, but rather is based on the formation of tertiary structure that allows specific recognition to take place.

259. Antisense molecules are designed to interact with a target nucleic acid molecule through either canonical or non-canonical base pairing. The interaction of the antisense molecule and the target molecule is designed to promote the destruction of the target molecule through, for example, RNAseH mediated RNA-DNA hybrid degradation. Alternatively the antisense molecule is designed to interrupt a processing function that normally would take place on the target molecule, such as transcription or replication. Antisense molecules can be designed based on the sequence of the target molecule. Numerous methods for optimization of antisense efficiency by finding the most accessible regions of the target molecule exist. Exemplary methods would be in vitro selection experiments and DNA modification studies using DMS and DEPC. It is preferred that antisense molecules bind the target molecule with a dissociation constant (k_d)less than 10^{-6}. It is more preferred that antisense molecules bind with a k_d less than 10^{-8}. It is also more preferred that the antisense molecules bind the target molecule with a k_d less than 10^{-10}. It is also preferred that the antisense molecules bind the target molecule with a k_d less than 10^{-12}. A representative sample of methods and techniques which aid in the design and use of antisense molecules can be found in the following non-limiting list of United States patents: 5,135,917, 5,294,533, 5,627,158, 5,641,754, 5,691,317, 5,780,607, 5,786,138, 5,849,903, 5,856,103, 5,919,772, 5,955,590, 5,990,088, 5,994,320, 5,998,602, 6,005,095, 6,007,995, 6,013,522, 6,017,898, 6,018,042, 6,025,198, 6,033,910, 6,040,296, 6,046,004, 6,046,319, and 6,057,437.

260. Aptamers are molecules that interact with a target molecule, preferably in a specific way. Typically aptamers are small nucleic acids ranging from 15-50 bases in length that fold into defined secondary and tertiary structures, such as stem-loops or G-quartets. Aptamers can bind small molecules, such as ATP (United States patent 5,631,146) and theophiline (United States patent 5,580,737), as well as large molecules, such as reverse transcriptase (United States patent 5,786,462) and thrombin (United States patent 5,543,293). Aptamers can bind very tightly with k_ds from the target molecule of less than 10^{-12} M. It is preferred that the aptamers bind the target molecule with a k_d less than 10^{-6}. It is more preferred that the aptamers bind the target molecule with a k_d less than 10^{-8}. It is also more preferred that the aptamers bind the target molecule with a k_d less than 10^{-10}. It is also preferred that the aptamers bind the target molecule with a k_d less than 10^{-12}. Aptamers can bind the target molecule with a very high degree of specificity. For example, aptamers have been isolated that have greater than a 10000 fold difference in binding affinities between the target molecule and another molecule that differ at only a single position on the molecule (United States patent 5,543,293). It is preferred that the aptamer have a k_d with the target molecule at least 10 fold lower than the k_d with a background binding molecule. It is more preferred that the aptamer have a k_d with the target molecule at least 100 fold lower than the k_d with a background binding molecule. It is more preferred that the aptamer have a k_d with the target molecule

at least 1000 fold lower than the k_d with a background binding molecule. It is preferred that the aptamer have a k_d with the target molecule at least 10000 fold lower than the k_d with a background binding molecule. It is preferred when doing the comparison for a polypeptide for example, that the background molecule be a different polypeptide. For example, when determining the specificity of TSG101 or Vps4A or Vps4B aptamers, such as TSG101 PTAP binding domain or UB binding domain aptamers, the background protein could be albumin. Representative examples of how to make and use aptamers to bind a variety of different target molecules can be found in the following non-limiting list of United States patents: 5,476,766, 5,503,978, 5,631,146, 5,731,424 , 5,780,228, 5,792,613, 5,795,721, 5,846,713, 5,858,660 , 5,861,254, 5,864,026, 5,869,641, 5,958,691, 6,001,988, 6,011,020, 6,013,443, 6,020,130, 6,028,186, 6,030,776, and 6,051,698.

261. Ribozymes are nucleic acid molecules that are capable of catalyzing a chemical reaction, either intramolecularly or intermolecularly. Ribozymes are thus catalytic nucleic acid. It is preferred that the ribozymes catalyze intermolecular reactions. There are a number of different types of ribozymes that catalyze nuclease or nucleic acid polymerase type reactions which are based on ribozymes found in natural systems, such as hammerhead ribozymes, (for example, but not limited to the following United States patents: 5,334,711, 5,436,330, 5,616,466, 5,633,133, 5,646,020, 5,652,094, 5,712,384, 5,770,715, 5,856,463, 5,861,288, 5,891,683, 5,891,684, 5,985,621, 5,989,908, 5,998,193, 5,998,203, WO 9858058 by Ludwig and Sproat, WO 9858057 by Ludwig and Sproat, and WO 9718312 by Ludwig and Sproat) hairpin ribozymes (for example, but not limited to the following United States patents: 5,631,115, 5,646,031, 5,683,902, 5,712,384, 5,856,188, 5,866,701, 5,869,339, and 6,022,962), and tetrahymena ribozymes (for example, but not limited to the following United States patents: 5,595,873 and 5,652,107). There are also a number of ribozymes that are not found in natural systems, but which have been engineered to catalyze specific reactions de novo (for example, but not limited to the following United States patents: 5,580,967, 5,688,670, 5,807,718, and 5,910,408). Preferred ribozymes cleave RNA or DNA substrates, and more preferably cleave RNA substrates. Ribozymes typically cleave nucleic acid substrates through recognition and binding of the target substrate with subsequent cleavage. This recognition is often based mostly on canonical or non-canonical base pair interactions. This property makes ribozymes particularly good candidates for target specific cleavage of nucleic acids because recognition of the target substrate is based on the target substrates sequence. Representative examples of how to make and use ribozymes to catalyze a variety of different reactions can be found in the following non-limiting list of United States patents: 5,646,042, 5,693,535, 5,731,295, 5,811,300, 5,837,855, 5,869,253, 5,877,021, 5,877,022, 5,972,699, 5,972,704, 5,989,906, and 6,017,756.

262. Triplex forming functional nucleic acid molecules are molecules that can interact with either double-stranded or single-stranded nucleic acid. When triplex molecules interact with a target region, a structure called a triplex is formed, in which there are three strands of DNA forming a complex dependant on both Watson-Crick and Hoogsteen base-pairing. Triplex molecules are preferred because they can bind target regions with high affinity and specificity. It is preferred that the triplex forming

molecules bind the target molecule with a k_d less than 10^{-6}. It is more preferred that the triplex forming molecules bind with a k_d less than 10^{-8}. It is also more preferred that the triplex forming molecules bind the target molecule with a k_d less than 10^{-10}. It is also preferred that the triplex forming molecules bind the target molecule with a k_d less than 10^{-12}. Representative examples of how to make and use triplex forming molecules to bind a variety of different target molecules can be found in the following non-limiting list of United States patents: 5,176,996, 5,645,985, 5,650,316, 5,683,874, 5,693,773, 5,834,185, 5,869,246, 5,874,566, and 5,962,426.

263. External guide sequences (EGSs) are molecules that bind a target nucleic acid molecule forming a complex, and this complex is recognized by RNase P, which cleaves the target molecule. EGSs can be designed to specifically target a RNA molecule of choice. RNAse P aids in processing transfer RNA (tRNA) within a cell. Bacterial RNAse P can be recruited to cleave virtually any RNA sequence by using an EGS that causes the target RNA:EGS complex to mimic the natural tRNA substrate. (WO 92/03566 by Yale, and Forster and Altman, Science 238:407-409 (1990)).

264. Similarly, eukaryotic EGS/RNAse P-directed cleavage of RNA can be utilized to cleave desired targets within eukaryotic cells. (Yuan et al., Proc. Natl. Acad. Sci. USA 89:8006-8010 (1992); WO 93/22434 by Yale; WO 95/24489 by Yale; Yuan and Altman, EMBO J 14:159-168 (1995), and Carrara et al., Proc. Natl. Acad. Sci. (USA) 92:2627-2631 (1995)). Representative examples of how to make and use EGS molecules to facilitate cleavage of a variety of different target molecules be found in the following non-limiting list of United States patents: 5,168,053, 5,624,824, 5,683,873, 5,728,521, 5,869,248, and 5,877,162

b) Antibodies

(1) Antibodies Generally

265. The term "antibodies" is used herein in a broad sense and includes both polyclonal and monoclonal antibodies. In addition to intact immunoglobulin molecules, also included in the term "antibodies" are fragments or polymers of those immunoglobulin molecules, and human or humanized versions of immunoglobulin molecules or fragments thereof, as long as they are chosen for their ability to interact with any of the proteins and compositions disclosed herein, such as, TSG101 or Vps4A or Vps4B or GAG, such that TSG101 or Vps4A or Vps4B or GAG is inhibited from promoting retroviral budding. Antibodies that bind the disclosed regions of any of the proteins and compositions disclosed herein, such as, TSG101 or Vps4A or Vps4B or GAG involved in the viral budding are also disclosed. The antibodies can be tested for their desired activity using the *in vitro* assays described herein, or by analogous methods, after which their *in vivo* therapeutic and/or prophylactic activities are tested according to known clinical testing methods. Also disclosed are functional equivalents of antibodies.

266. The term "monoclonal antibody" as used herein refers to an antibody obtained from a substantially homogeneous population of antibodies, i.e., the individual antibodies within the population are identical except for possible naturally occurring mutations that may be present in a small subset of the antibody molecules. The monoclonal antibodies herein specifically include "chimeric" antibodies in

which a portion of the heavy and/or light chain is identical with or homologous to corresponding sequences in antibodies derived from a particular species or belonging to a particular antibody class or subclass, while the remainder of the chain(s) is identical with or homologous to corresponding sequences in antibodies derived from another species or belonging to another antibody class or subclass, as well as fragments of such antibodies, as long as they exhibit the desired antagonistic activity (See, U.S. Pat. No. 4,816,567 and Morrison et al., *Proc. Natl. Acad. Sci. USA*, 81:6851-6855 (1984)).

267. The disclosed monoclonal antibodies can be made using any procedure which produces mono clonal antibodies. For example, monoclonal antibodies of the invention can be prepared using hybridoma methods, such as those described by Kohler and Milstein, *Nature*, 256:495 (1975). In a hybridoma method, a mouse or other appropriate host animal is typically immunized with an immunizing agent to elicit lymphocytes that produce or are capable of producing antibodies that will specifically bind to the immunizing agent.

268. The monoclonal antibodies may also be made by recombinant DNA methods, such as those described in U.S. Pat. No. 4,816,567 (Cabilly et al.). DNA encoding the disclosed monoclonal antibodies can be readily isolated and sequenced using conventional procedures (e.g., by using oligonucleotide probes that are capable of binding specifically to genes encoding the heavy and light chains of murine antibodies). Libraries of antibodies or active antibody fragments can also be generated and screened using phage display techniques, e.g., as described in U.S. Patent No. 5,804,440 to Burton et al. and U.S. Patent No. 6,096,441 to Barbas et al.

269. *In vitro* methods are also suitable for preparing monovalent antibodies. Digestion of antibodies to produce fragments thereof, particularly, Fab fragments, can be accomplished using routine techniques known in the art. For instance, digestion can be performed using papain. Examples of papain digestion are described in WO 94/29348 published Dec. 22, 1994 and U.S. Pat. No. 4,342,566. Papain digestion of antibodies typically produces two identical antigen binding fragments, called Fab fragments, each with a single antigen binding site, and a residual Fc fragment. Pepsin treatment yields a fragment that has two antigen combining sites and is still capable of cross-linking antigen.

270. The fragments, whether attached to other sequences or not, can also include insertions, deletions, substitutions, or other selected modifications of particular regions or specific amino acids residues, provided the activity of the antibody or antibody fragment is not significantly altered or impaired compared to the non-modified antibody or antibody fragment. These modifications can provide for some additional property, such as to remove/add amino acids capable of disulfide bonding, to increase its bio-longevity, to alter its secretory characteristics, etc. In any case, the antibody or antibody fragment must possess a bioactive property, such as specific binding to its cognate antigen. Functional or active regions of the antibody or antibody fragment may be identified by mutagenesis of a specific region of the protein, followed by expression and testing of the expressed polypeptide. Such methods are readily apparent to a skilled practitioner in the art and can include site-specific mutagenesis of the nucleic acid encoding the antibody or antibody fragment. (Zoller, M.J. *Curr. Opin. Biotechnol.* 3:348-354, 1992).

271. As used herein, the term "antibody" or "antibodies" can also refer to a human antibody and/or a humanized antibody. Many non-human antibodies (e.g., those derived from mice, rats, or rabbits) are naturally antigenic in humans, and thus can give rise to undesirable immune responses when administered to humans. Therefore, the use of human or humanized antibodies in the methods of the invention serves to lessen the chance that an antibody administered to a human will evoke an undesirable immune response.

(2) Human antibodies

272. The human antibodies of the invention can be prepared using any technique. Examples of techniques for human monoclonal antibody production include those described by Cole et al. (*Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, p. 77, 1985) and by Boerner et al. (*J. Immunol.*, 147(1):86-95, 1991). Human antibodies of the invention (and fragments thereof) can also be produced using phage display libraries (Hoogenboom et al., *J. Mol. Biol.*, 227:381, 1991; Marks et al., *J. Mol. Biol.*, 222:581, 1991).

273. The human antibodies of the invention can also be obtained from transgenic animals. For example, transgenic, mutant mice that are capable of producing a full repertoire of human antibodies, in response to immunization, have been described (see, e.g., Jakobovits et al., *Proc. Natl. Acad. Sci. USA*, 90:2551-255 (1993); Jakobovits et al., *Nature*, 362:255-258 (1993); Bruggermann et al., *Year in Immunol.*, 7:33 (1993)). Specifically, the homozygous deletion of the antibody heavy chain joining region (*J(H)*) gene in these chimeric and germ-line mutant mice results in complete inhibition of endogenous antibody production, and the successful transfer of the human germ-line antibody gene array into such germ-line mutant mice results in the production of human antibodies upon antigen challenge.

(3) Humanized antibodies

274. Antibody humanization techniques generally involve the use of recombinant DNA technology to manipulate the DNA sequence encoding one or more polypeptide chains of an antibody molecule. Accordingly, a humanized form of a non-human antibody (or a fragment thereof) is a chimeric antibody or antibody chain (or a fragment thereof, such as an Fv, Fab, Fab', or other antigen-binding portion of an antibody) which contains a portion of an antigen binding site from a non-human (donor) antibody integrated into the framework of a human (recipient) antibody.

275. To generate a humanized antibody, residues from one or more complementarity determining regions (CDRs) of a recipient (human) antibody molecule are replaced by residues from one or more CDRs of a donor (non-human) antibody molecule that is known to have desired antigen binding characteristics (e.g., a certain level of specificity and affinity for the target antigen). In some instances, Fv framework (FR) residues of the human antibody are replaced by corresponding non-human residues. Humanized antibodies may also contain residues which are found neither in the recipient antibody nor in the imported CDR or framework sequences. Generally, a humanized antibody has one or more amino acid residues introduced into it from a source which is non-human. In practice, humanized antibodies are typically human antibodies in which some CDR residues and possibly some FR residues are substituted

by residues from analogous sites in rodent antibodies. Humanized antibodies generally contain at least a portion of an antibody constant region (Fc), typically that of a human antibody (Jones et al., *Nature*, 321:522-525 (1986), Reichmann et al., *Nature*, 332:323-327 (1988), and Presta, *Curr. Opin. Struct. Biol.*, 2:593-596 (1992)).

276. Methods for humanizing non-human antibodies are well known in the art. For example, humanized antibodies can be generated according to the methods of Winter and co-workers (Jones et al., *Nature*, 321:522-525 (1986), Riechmann et al., *Nature*, 332:323-327 (1988), Verhoeyen et al., *Science*, 239:1534-1536 (1988)), by substituting rodent CDRs or CDR sequences for the corresponding sequences of a human antibody. Methods that can be used to produce humanized antibodies are also described in U.S. Patent No. 4,816,567 (Cabilly et al.), U.S. Patent No. 5,565,332 (Hoogenboom et al.), U.S. Patent No. 5,721,367 (Kay et al.), U.S. Patent No. 5,837,243 (Deo et al.), U.S. Patent No. 5, 939,598 (Kucherlapati et al.), U.S. Patent No. 6,130,364 (Jakobovits et al.), and U.S. Patent No. 6,180,377 (Morgan et al.).

(4) Administration of antibodies

277. Administration of the antibodies can be done as disclosed herein, as the delivery of any non-nucleic acid composition or pharmaceutical. Nucleic acid approaches for antibody delivery also exist. The broadly neutralizing anti TSG101 or Vps4A or Vps4B or GAG antibodies and antibody fragments of the invention can also be administered to patients or subjects as a nucleic acid preparation (e.g., DNA or RNA) that encodes the antibody or antibody fragment, such that the patient's or subject's own cells take up the nucleic acid and produce and secrete the encoded antibody or antibody fragment. The delivery of the nucleic acid can be by any means, as disclosed herein, for example.

c) Molecules that impair ATPase function of Vps4A and Vps4B

278. Disclosed are variants of Vps4A and Vps4B that have impaired ATPase activity or which themselves act as negative regulators of ATPase activity of Vps4A and Vps4B. These molecules when administered to cells harboring a virus such as HIV or ebola can inhibit the budding of the virus. Thus, molecules or mutants that inhibit ATPase activity are preferred and disclosed. One example of such a mutant is the ATPase-defective GFP-Vps4 fusion proteins induces formation of enlarged endosomes (called "class E" compartments in yeast) that are defective in the sorting and recycling of endocytosed substrates (Bishop and Woodman, 2000) (which are herein incorporated by reference for material at least related to Vps4 and mutants thereof). Other mutants, for example, are two different dominant negative mutations: K173Q, which blocks ATP binding ($Vps4_{K173Q}$), and E228Q, which blocks ATP hydrolysis ($Vps4_{E228Q}$) at the positions noted in Vps4. This shows that ATPase activity of Vps4 is required for retroviral, such as HIV budding, and that any mutation or molecule that inhibits Vps4 ATPase activity will function as and is disclosed herein. Mutants such as this exist for both Vps4A and Vps4B.

d) Compositions identified by screening with disclosed compositions / combinatorial chemistry or by using the structure information

(1) Combinatorial chemistry

279. The disclosed compositions can be used as targets for any combinatorial technique to identify molecules or macromolecular molecules that interact with the disclosed compositions in a desired way. The nucleic acids, peptides, and related molecules disclosed herein can be used as targets for the combinatorial approaches. Also disclosed are the compositions that are identified through combinatorial techniques or screening techniques, or through computer modeling techniques, in which the compositions herein disclosed and their structural information, or portions thereof, are used as the target in a combinatorial or screening protocol, or can direct combinatorial or screening protocols.

280. It is understood that when using the disclosed compositions in combinatorial techniques or screening methods, molecules, such as macromolecular molecules, will be identified that have particular desired properties such as inhibition or stimulation or the target molecule's function. The molecules identified and isolated when using the disclosed compositions, such as, any of the proteins disclosed herein, such asTSG101 or Vps4A or Vps4B or GAG, are also disclosed. Thus, the products produced using the combinatorial or screening approaches that involve the disclosed compositions, such as any of the proteins disclosed herein, such as TSG101 or Vps4A or Vps4B or GAG, are also considered herein disclosed.

281. Combinatorial chemistry includes but is not limited to all methods for isolating small molecules or macromolecules that are capable of binding either a small molecule or another macromolecule, typically in an iterative process. Proteins, oligonucleotides, and sugars are examples of macromolecules. For example, oligonucleotide molecules with a given function, catalytic or ligand-binding, can be isolated from a complex mixture of random oligonucleotides in what has been referred to as "in vitro genetics" (Szostak, TIBS 19:89, 1992). One synthesizes a large pool of molecules bearing random and defined sequences and subjects that complex mixture, for example, approximately 10^{15} individual sequences in 100 μg of a 100 nucleotide RNA, to some selection and enrichment process. Through repeated cycles of affinity chromatography and PCR amplification of the molecules bound to the ligand on the column, Ellington and Szostak (1990) estimated that 1 in 10^{10} RNA molecules folded in such a way as to bind a small molecule dyes. DNA molecules with such ligand-binding behavior have been isolated as well (Ellington and Szostak, 1992; Bock et al, 1992). Techniques aimed at similar goals exist for small organic molecules, proteins, antibodies and other macromolecules known to those of skill in the art. Screening sets of molecules for a desired activity whether based on small organic libraries, oligonucleotides, or antibodies is broadly referred to as combinatorial chemistry. Combinatorial techniques are particularly suited for defining binding interactions between molecules and for isolating molecules that have a specific binding activity, often called aptamers when the macromolecules are nucleic acids.

282. There are a number of methods for isolating proteins which either have de novo activity or a modified activity. For example, phage display libraries have been used to isolate numerous peptides that interact with a specific target. (See for example, United States Patent No. 6,031,071; 5,824,520; 5,596,079; and 5,565,332 which are herein incorporated by reference at least for their material related to phage display and methods relate to combinatorial chemistry)

283. A preferred method for isolating proteins that have a given function is described by Roberts and Szostak (Roberts R.W. and Szostak J.W. Proc. Natl. Acad. Sci. USA, 94(23)12997-302 (1997). This combinatorial chemistry method couples the functional power of proteins and the genetic power of nucleic acids. An RNA molecule is generated in which a puromycin molecule is covalently attached to the 3'-end of the RNA molecule. An *in vitro* translation of this modified RNA molecule causes the correct protein, encoded by the RNA to be translated. In addition, because of the attachment of the puromycin, a peptidyl acceptor which cannot be extended, the growing peptide chain is attached to the puromycin which is attached to the RNA. Thus, the protein molecule is attached to the genetic material that encodes it. Normal *in vitro* selection procedures can now be done to isolate functional peptides. Once the selection procedure for peptide function is complete traditional nucleic acid manipulation procedures are performed to amplify the nucleic acid that codes for the selected functional peptides. After amplification of the genetic material, new RNA is transcribed with puromycin at the 3'-end, new peptide is translated and another functional round of selection is performed. Thus, protein selection can be performed in an iterative manner just like nucleic acid selection techniques. The peptide which is translated is controlled by the sequence of the RNA attached to the puromycin. This sequence can be anything from a random sequence engineered for optimum translation (i.e. no stop codons etc.) or it can be a degenerate sequence of a known RNA molecule to look for improved or altered function of a known peptide. The conditions for nucleic acid amplification and in vitro translation are well known to those of ordinary skill in the art and are preferably performed as in Roberts and Szostak (Roberts R.W. and Szostak J.W. Proc. Natl. Acad. Sci. USA, 94(23)12997-302 (1997)).

284. Another preferred method for combinatorial methods designed to isolate peptides is described in Cohen et al. (Cohen B.A.,et al., Proc. Natl. Acad. Sci. USA 95(24):14272-7 (1998)). This method utilizes and modifies two-hybrid technology. Yeast two-hybrid systems are useful for the detection and analysis of protein:protein interactions. The two-hybrid system, initially described in the yeast *Saccharomyces cerevisiae*, is a powerful molecular genetic technique for identifying new regulatory molecules, specific to the protein of interest (Fields and Song, *Nature* 340:245-6 (1989)). Cohen et al., modified this technology so that novel interactions between synthetic or engineered peptide sequences could be identified which bind a molecule of choice. The benefit of this type of technology is that the selection is done in an intracellular environment. The method utilizes a library of peptide molecules that attached to an acidic activation domain. A peptide of choice, for example, a portion of TSG101 or Vps4A or Vps4B or GAG, such as the TSG101 PTAP binding domain, is attached to a DNA binding domain of a transcriptional activation protein, such as Gal 4. By performing the two-hybrid

technique on this type of system, molecules that bind the TSG101 or Vps4A or Vps4B or GAG can be identified.

285. Using methodology well known to those of skill in the art, in combination with various combinatorial libraries, one can isolate and characterize those small molecules or macromolecules, which bind to or interact with the desired target. The relative binding affinity of these compounds can be compared and optimum compounds identified using competitive binding studies, which are well known to those of skill in the art.

286. Techniques for making combinatorial libraries and screening combinatorial libraries to isolate molecules which bind a desired target are well known to those of skill in the art. Representative techniques and methods can be found in but are not limited to United States patents 5,084,824, 5,288,514, 5,449,754, 5,506,337, 5,539,083, 5,545,568, 5,556,762, 5,565,324, 5,565,332, 5,573,905, 5,618,825, 5,619,680, 5,627,210, 5,646,285, 5,663,046, 5,670,326, 5,677,195, 5,683,899, 5,688,696, 5,688,997, 5,698,685, 5,712,146, 5,721,099, 5,723,598, 5,741,713, 5,792,431, 5,807,683, 5,807,754, 5,821,130, 5,831,014, 5,834,195, 5,834,318, 5,834,588, 5,840,500, 5,847,150, 5,856,107, 5,856,496, 5,859,190, 5,864,010, 5,874,443, 5,877,214, 5,880,972, 5,886,126, 5,886,127, 5,891,737, 5,916,899, 5,919,955, 5,925,527, 5,939,268, 5,942,387, 5,945,070, 5,948,696, 5,958,702, 5,958,792, 5,962,337, 5,965,719, 5,972,719, 5,976,894, 5,980,704, 5,985,356, 5,999,086, 6,001,579, 6,004,617, 6,008,321, 6,017,768, 6,025,371, 6,030,917, 6,040,193, 6,045,671, 6,045,755, 6,060,596, and 6,061,636.

287. Combinatorial libraries can be made from a wide array of molecules using a number of different synthetic techniques. For example, libraries containing fused 2,4-pyrimidinediones (United States patent 6,025,371) dihydrobenzopyrans (United States Patent 6,017,768and 5,821,130), amide alcohols (United States Patent 5,976,894), hydroxy-amino acid amides (United States Patent 5,972,719) carbohydrates (United States patent 5,965,719), 1,4-benzodiazepin-2,5-diones (United States patent 5,962,337), cyclics (United States patent 5,958,792), biaryl amino acid amides (United States patent 5,948,696), thiophenes (United States patent 5,942,387), tricyclic Tetrahydroquinolines (United States patent 5,925,527), benzofurans (United States patent 5,919,955), isoquinolines (United States patent 5,916,899), hydantoin and thiohydantoin (United States patent 5,859,190), indoles (United States patent 5,856,496), imidazol-pyrido-indole and imidazol-pyrido-benzothiophenes (United States patent 5,856,107) substituted 2-methylene-2, 3-dihydrothiazoles (United States patent 5,847,150), quinolines (United States patent 5,840,500), PNA (United States patent 5,831,014), containing tags (United States patent 5,721,099), polyketides (United States patent 5,712,146), morpholino-subunits (United States patent 5,698,685 and 5,506,337), sulfamides (United States patent 5,618,825), and benzodiazepines (United States patent 5,288,514).

288. As used herein combinatorial methods and libraries included traditional screening methods and libraries as well as methods and libraries used in iterative processes.

5. Computer readable mediums

289. It is understood that the disclosed nucleic acids and proteins can be represented as a sequence consisting of the nucleotides of amino acids. There are a variety of ways to display these sequences, for example the nucleotide guanosine can be represented by G or g. Likewise the amino acid valine can be represented by Val or V. Those of skill in the art understand how to display and express any nucleic acid or protein sequence in any of the variety of ways that exist, each of which is considered herein disclosed. Specifically contemplated herein is the display of these sequences on computer readable mediums, such as, commercially available floppy disks, tapes, chips, hard drives, compact disks, and video disks, or other computer readable mediums. Also disclosed are the binary code representations of the disclosed sequences. Those of skill in the art understand what computer readable mediums. Thus, computer readable mediums on which the nucleic acids or protein sequences are recorded, stored, or saved.

6. General composition information

a) Sequence similarities

290. It is understood that as discussed herein the use of the terms homology and identity mean the same thing as similarity. Thus, for example, if the use of the word homology is used between two non-natural sequences it is understood that this is not necessarily indicating an evolutionary relationship between these two sequences, but rather is looking at the similarity or relatedness between their nucleic acid sequences. Many of the methods for determining homology between two evolutionarily related molecules are routinely applied to any two or more nucleic acids or proteins for the purpose of measuring sequence similarity regardless of whether they are evolutionarily related or not.

291. In general, it is understood that one way to define any known variants and derivatives or those that might arise, of the disclosed genes and proteins herein, is through defining the variants and derivatives in terms of homology to specific known sequences. This identity of particular sequences disclosed herein is also discussed elsewhere herein. In general, variants of genes and proteins herein disclosed typically have at least, about 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, or 99 percent homology to the stated sequence or the native sequence. Those of skill in the art readily understand how to determine the homology of two proteins or nucleic acids, such as genes. For example, the homology can be calculated after aligning the two sequences so that the homology is at its highest level.

292. Another way of calculating homology can be performed by published algorithms. Optimal alignment of sequences for comparison may be conducted by the local homology algorithm of Smith and Waterman Adv. Appl. Math. 2: 482 (1981), by the homology alignment algorithm of Needleman and Wunsch, J. MoL Biol. 48: 443 (1970), by the search for similarity method of Pearson and Lipman, Proc. Natl. Acad. Sci. U.S.A. 85: 2444 (1988), by computerized implementations of these algorithms (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, WI), or by inspection.

293. The same types of homology can be obtained for nucleic acids by for example the algorithms disclosed in Zuker, M. *Science* 244:48-52, 1989, Jaeger et al. *Proc. Natl. Acad. Sci. USA* 86:7706-7710, 1989, Jaeger et al. *Methods Enzymol.* 183:281-306, 1989 which are herein incorporated by reference for at least material related to nucleic acid alignment. It is understood that any of the methods typically can be used and that in certain instances the results of these various methods may differ, but the skilled artisan understands if identity is found with at least one of these methods, the sequences would be said to have the stated identity, and be disclosed herein.

294. For example, as used herein, a sequence recited as having a particular percent homology to another sequence refers to sequences that have the recited homology as calculated by any one or more of the calculation methods described above. For example, a first sequence has 80 percent homology, as defined herein, to a second sequence if the first sequence is calculated to have 80 percent homology to the second sequence using the Zuker calculation method even if the first sequence does not have 80 percent homology to the second sequence as calculated by any of the other calculation methods. As another example, a first sequence has 80 percent homology, as defined herein, to a second sequence if the first sequence is calculated to have 80 percent homology to the second sequence using both the Zuker calculation method and the Pearson and Lipman calculation method even if the first sequence does not have 80 percent homology to the second sequence as calculated by the Smith and Waterman calculation method, the Needleman and Wunsch calculation method, the Jaeger calculation methods, or any of the other calculation methods. As yet another example, a first sequence has 80 percent homology, as defined herein, to a second sequence if the first sequence is calculated to have 80 percent homology to the second sequence using each of calculation methods (although, in practice, the different calculation methods will often result in different calculated homology percentages).

b) Hybridization

295. The term hybridization typically means a sequence driven interaction between at least two nucleic acid molecules, such as a primer or a probe and a gene. Sequence driven interaction means an interaction that occurs between two nucleotides or nucleotide analogs or nucleotide derivatives in a nucleotide specific manner. For example, G interacting with C or A interacting with T are sequence driven interactions. Typically sequence driven interactions occur on the Watson-Crick face or Hoogsteen face of the nucleotide. The hybridization of two nucleic acids is affected by a number of conditions and parameters known to those of skill in the art. For example, the salt concentrations, pH, and temperature of the reaction all affect whether two nucleic acid molecules will hybridize.

296. Parameters for selective hybridization between two nucleic acid molecules are well known to those of skill in the art. For example, in some embodiments selective hybridization conditions can be defined as stringent hybridization conditions. For example, stringency of hybridization is controlled by both temperature and salt concentration of either or both of the hybridization and washing steps. For example, the conditions of hybridization to achieve selective hybridization may involve hybridization in high ionic strength solution (6X SSC or 6X SSPE) at a temperature that is about 12-25°C below the Tm

(the melting temperature at which half of the molecules dissociate from their hybridization partners) followed by washing at a combination of temperature and salt concentration chosen so that the washing temperature is about 5°C to 20°C below the Tm. The temperature and salt conditions are readily determined empirically in preliminary experiments in which samples of reference DNA immobilized on filters are hybridized to a labeled nucleic acid of interest and then washed under conditions of different stringencies. Hybridization temperatures are typically higher for DNA-RNA and RNA-RNA hybridizations. The conditions can be used as described above to achieve stringency, or as is known in the art. (Sambrook et al., Molecular Cloning: A Laboratory Manual, 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 1989; Kunkel et al. Methods Enzymol. 1987:154:367, 1987 which is herein incorporated by reference for material at least related to hybridization of nucleic acids). A preferable stringent hybridization condition for a DNA:DNA hybridization can be at about 68°C (in aqueous solution) in 6X SSC or 6X SSPE followed by washing at 68°C. Stringency of hybridization and washing, if desired, can be reduced accordingly as the degree of complementarity desired is decreased, and further, depending upon the G-C or A-T richness of any area wherein variability is searched for. Likewise, stringency of hybridization and washing, if desired, can be increased accordingly as homology desired is increased, and further, depending upon the G-C or A-T richness of any area wherein high homology is desired, all as known in the art.

297. Another way to define selective hybridization is by looking at the amount (percentage) of one of the nucleic acids bound to the other nucleic acid. For example, in some embodiments selective hybridization conditions would be when at least about, 60, 65, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100 percent of the limiting nucleic acid is bound to the non-limiting nucleic acid. Typically, the non-limiting primer is in for example, 10 or 100 or 1000 fold excess. This type of assay can be performed at under conditions where both the limiting and non-limiting primer are for example, 10 fold or 100 fold or 1000 fold below their k_d, or where only one of the nucleic acid molecules is 10 fold or 100 fold or 1000 fold or where one or both nucleic acid molecules are above their k_d.

298. Another way to define selective hybridization is by looking at the percentage of primer that gets enzymatically manipulated under conditions where hybridization is required to promote the desired enzymatic manipulation. For example, in some embodiments selective hybridization conditions would be when at least about, 60, 65, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100 percent of the primer is enzymatically manipulated under conditions which promote the enzymatic manipulation, for example if the enzymatic manipulation is DNA extension, then selective hybridization conditions would be when at least about 60, 65, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100 percent of the primer molecules are extended. Preferred conditions also include those suggested by the manufacturer or indicated in the art as being appropriate for the enzyme performing the manipulation.

299. Just as with homology, it is understood that there are a variety of methods herein disclosed for determining the level of hybridization between two nucleic acid molecules. It is understood that these methods and conditions may provide different percentages of hybridization between two nucleic acid molecules, but unless otherwise indicated meeting the parameters of any of the methods would be sufficient. For example if 80% hybridization was required and as long as hybridization occurs within the required parameters in any one of these methods it is considered disclosed herein.

300. It is understood that those of skill in the art understand that if a composition or method meets any one of these criteria for determining hybridization either collectively or singly it is a composition or method that is disclosed herein.

c) Nucleic acids

301. There are a variety of molecules disclosed herein that are nucleic acid based, including for example the nucleic acids that encode, for example TSG101, Vps4A, or Vps4B, Bc2, Vps28, or Vps37, as well as various functional nucleic acids. The disclosed nucleic acids are made up of, for example, nucleotides, nucleotide analogs, or nucleotide substitutes. Non-limiting examples of these and other molecules are discussed herein. It is understood that for example, when a vector is expressed in a cell, that the expressed mRNA will typically be made up of A, C, G, and U. Likewise, it is understood that if, for example, an antisense molecule is introduced into a cell or cell environment through for example exogenous delivery, it is advantagous that the antisense molecule be made up of nucleotide analogs that reduce the degradation of the antisense molecule in the cellular environment.

(1) Nucleotides and related molecules

302. A nucleotide is a molecule that contains a base moiety, a sugar moiety and a phosphate moiety. Nucleotides can be linked together through their phosphate moieties and sugar moieties creating an internucleoside linkage. The base moiety of a nucleotide can be adenin-9-yl (A), cytosin-1-yl (C), guanin-9-yl (G), uracil-1-yl (U), and thymin-1-yl (T). The sugar moiety of a nucleotide is a ribose or a deoxyribose. The phosphate moiety of a nucleotide is pentavalent phosphate. An non-limiting example of a nucleotide would be 3'-AMP (3'-adenosine monophosphate) or 5'-GMP (5'-guanosine monophosphate).

303. A nucleotide analog is a nucleotide which contains some type of modification to either the base, sugar, or phosphate moieties. Modifications to nucleotides are well known in the art and would include for example, 5-methylcytosine (5-me-C), 5-hydroxymethyl cytosine, xanthine, hypoxanthine, and 2-aminoadenine as well as modifications at the sugar or phosphate moieties.

304. Nucleotide substitutes are molecules having similar functional properties to nucleotides, but which do not contain a phosphate moiety, such as peptide nucleic acid (PNA). Nucleotide substitutes are molecules that will recognize nucleic acids in a Watson-Crick or Hoogsteen manner, but which are linked together through a moiety other than a phosphate moiety. Nucleotide substitutes are able to conform to a double helix type structure when interacting with the appropriate target nucleic acid.

305. It is also possible to link other types of molecules (conjugates) to nucleotides or nucleotide analogs to enhance for example, cellular uptake. Conjugates can be chemically linked to the nucleotide or nucleotide analogs. Such conjugates include but are not limited to lipid moieties such as a cholesterol moiety. (Letsinger et al., Proc. Natl. Acad. Sci. USA, 1989,86, 6553-6556),

306. A Watson-Crick interaction is at least one interaction with the Watson-Crick face of a nucleotide, nucleotide analog, or nucleotide substitute. The Watson-Crick face of a nucleotide, nucleotide analog, or nucleotide substitute includes the C2, N1, and C6 positions of a purine based nucleotide, nucleotide analog, or nucleotide substitute and the C2, N3, C4 positions of a pyrimidine based nucleotide, nucleotide analog, or nucleotide substitute.

307. A Hoogsteen interaction is the interaction that takes place on the Hoogsteen face of a nucleotide or nucleotide analog, which is exposed in the major groove of duplex DNA. The Hoogsteen face includes the N7 position and reactive groups (NH2 or O) at the C6 position of purine nucleotides.

(2) Sequences

308. There are a variety of sequences related to the TSG101 (See for example, Q99816, Q61187 NP_006283, NP_057292, NP_068684), Vps4A, or Vps4B, Bc2, Vps37, Vps28, and GAG, these sequences and others are herein incorporated by reference in their entireties as well as for individual subsequences contained therein.

309. It is understood that the description related to the disclosed sequences is applicable to any sequence related to any of the proteins and genes disclosed herein, such as TSG101, Vps4A, or Vps4B or retroviral GAG, for example, unless specifically indicated otherwise. Those of skill in the art understand how to resolve sequence discrepancies and differences and to adjust the compositions and methods relating to a particular sequence to other related sequences (i.e. sequences of TSG101, Vps4A, or Vps4B). Primers and/or probes can be designed for any of the disclosed molecules, such as TSG101, Vps4A, or Vps4B sequence given the information disclosed herein and known in the art.

(3) Primers and probes

310. Disclosed are compositions including primers and probes, which are capable of interacting with any of the disclosed nucleic acids, such as the TSG101, Vps4A, or Vps4B genes or GAG as disclosed herein. In certain embodiments the primers are used to support DNA amplification reactions. Typically the primers will be capable of being extended in a sequence specific manner. Extension of a primer in a sequence specific manner includes any methods wherein the sequence and/or composition of the nucleic acid molecule to which the primer is hybridized or otherwise associated directs or influences the composition or sequence of the product produced by the extension of the primer. Extension of the primer in a sequence specific manner therefore includes, but is not limited to, PCR, DNA sequencing, DNA extension, DNA polymerization, RNA transcription, or reverse transcription. Techniques and conditions that amplify the primer in a sequence specific manner are preferred. In certain embodiments the primers are used for the DNA amplification reactions, such as PCR or direct sequencing. It is understood that in certain embodiments the primers can also be extended using non-enzymatic

techniques, where for example, the nucleotides or oligonucleotides used to extend the primer are modified such that they will chemically react to extend the primer in a sequence specific manner. Typically the disclosed primers hybridize with any of the disclosed nucleic acids, such as TSG101, Vps4A, or Vps4B nucleic acid or region of the TSG101, Vps4A, or Vps4B nucleic acid or they hybridize with the complement of the TSG101, Vps4A, or Vps4B nucleic acid or complement of a region of the TSG101, Vps4A, or Vps4B nucleic acid.

d) Delivery of the compositions to cells

(1) Nucleic Acid Delivery

311. There are a number of compositions and methods which can be used to deliver nucleic acids to cells, either in vitro or in vivo. These methods and compositions can largely be broken down into two classes: viral based delivery systems and non-viral based delivery systems. For example, the nucleic acids can be delivered through a number of direct delivery systems such as, electroporation, lipofection, calcium phosphate precipitation, plasmids, viral vectors, viral nucleic acids, phage nucleic acids, phages, cosmids, or via transfer of genetic material in cells or carriers such as cationic liposomes. Appropriate means for transfection, including viral vectors, chemical transfectants, or physico-mechanical methods such as electroporation and direct diffusion of DNA, are described by, for example, Wolff, J. A., et al., Science, 247, 1465-1468, (1990); and Wolff, J. A. Nature, 352, 815-818, (1991) Such methods are well known in the art and readily adaptable for use with the compositions and methods described herein. In certain cases, the methods will be modified to specifically function with large DNA molecules. Further, these methods can be used to target certain diseases and cell populations by using the targeting characteristics of the carrier.

312. In the methods described herein, which include the administration and uptake of exogenous DNA into the cells of a subject (i.e., gene transduction or transfection), the nucleic acids of the present invention can be in the form of naked DNA or RNA, or the nucleic acids can be in a vector for delivering the nucleic acids to the cells, whereby the encoding DNA or DNA or fragment is under the transcriptional regulation of a promoter, as would be well understood by one of ordinary skill in the art as well as enhancers. The vector can be a commercially available preparation, such as an adenovirus vector (Quantum Biotechnologies, Inc. (Laval, Quebec, Canada).

313. As one example, vector delivery can be via a viral system, such as a retroviral vector system which can package a recombinant retroviral genome (see e.g., Pastan et al., *Proc. Natl. Acad. Sci. U.S.A.* 85:4486, 1988; Miller et al., *Mol. Cell. Biol.* 6:2895, 1986). The recombinant retrovirus can then be used to infect and thereby deliver to the infected cells nucleic acid encoding a broadly neutralizing antibody (or active fragment thereof) of the invention. The exact method of introducing the altered nucleic acid into mammalian cells is, of course, not limited to the use of retroviral vectors. Other techniques are widely available for this procedure including the use of adenoviral vectors (Mitani et al., *Hum. Gene Ther.* 5:941-948, 1994), adeno-associated viral (AAV) vectors (Goodman et al., *Blood* 84:1492-1500, 1994), lentiviral vectors (Naidini et al., *Science* 272:263-267, 1996), pseudotyped retroviral vectors

(Agrawal et al., *Exper. Hematol.* 24:738-747, 1996). Physical transduction techniques can also be used, such as liposome delivery and receptor-mediated and other endocytosis mechanisms (see, for example, Schwartzenberger et al., *Blood* 87:472-478, 1996). This invention can be used in conjunction with any of these or other commonly used gene transfer methods.

314. As one example, if the antibody-encoding nucleic acid or some other nucleic acid encoding an inhibitor of the TSG101, Vps4A, or Vps4B-GAG interactions of the invention is delivered to the cells of a subject in an adenovirus vector, the dosage for administration of adenovirus to humans can range from about 10^7 to 10^9 plaque forming units (pfu) per injection but can be as high as 10^{12} pfu per injection (Crystal, *Hum. Gene Ther.* 8:985-1001, 1997; Alvarez and Curiel, *Hum. Gene Ther.* 8:597-613, 1997). A subject can receive a single injection, or, if additional injections are necessary, they can be repeated at six month intervals (or other appropriate time intervals, as determined by the skilled practitioner) for an indefinite period and/or until the efficacy of the treatment has been established.

315. Parenteral administration of the nucleic acid or vector of the present invention, if used, is generally characterized by injection. Injectables can be prepared in conventional forms, either as liquid solutions or suspensions, solid forms suitable for solution of suspension in liquid prior to injection, or as emulsions. A more recently revised approach for parenteral administration involves use of a slow release or sustained release system such that a constant dosage is maintained. See, e.g., U.S. Patent No. 3,610,795, which is incorporated by reference herein. For additional discussion of suitable formulations and various routes of administration of therapeutic compounds, see, e.g., *Remington: The Science and Practice of Pharmacy* (19th ed.) ed. A.R. Gennaro, Mack Publishing Company, Easton, PA 1995.

316. Nucleic acids that are delivered to cells which are to be integrated into the host cell genome, typically contain integration sequences. These sequences are often viral related sequences, particularly when viral based systems are used. These viral integration systems can also be incorporated into nucleic acids which are to be delivered using a non-nucleic acid based system of deliver, such as a liposome, so that the nucleic acid contained in the delivery system can be come integrated into the host genome.

317. Other general techniques for integration into the host genome include, for example, systems designed to promote homologous recombination with the host genome. These systems typically rely on sequence flanking the nucleic acid to be expressed that has enough homology with a target sequence within the host cell genome that recombination between the vector nucleic acid and the target nucleic acid takes place, causing the delivered nucleic acid to be integrated into the host genome. These systems and the methods necessary to promote homologous recombination are known to those of skill in the art.

(2) Non-nucleic acid based systems

318. The disclosed compositions can be delivered to the target cells in a variety of ways. For example, the compositions can be delivered through electroporation, or through lipofection, or through calcium phosphate precipitation. The delivery mechanism chosen will depend in part on the type of cell targeted and whether the delivery is occurring for example in vivo or in vitro.

319. Thus, the compositions can comprise, in addition to the disclosed compositions or vectors for example, lipids such as liposomes, such as cationic liposomes (e.g., DOTMA, DOPE, DC-cholesterol) or anionic liposomes. Liposomes can further comprise proteins to facilitate targeting a particular cell, if desired. Administration of a composition comprising a compound and a cationic liposome can be administered to the blood afferent to a target organ or inhaled into the respiratory tract to target cells of the respiratory tract. Regarding liposomes, see, e.g., Brigham et al. *Am. J. Resp. Cell. Mol. Biol.* 1:95-100 (1989); Felgner et al. *Proc. Natl. Acad. Sci USA* 84:7413-7417 (1987); U.S. Pat. No.4,897,355. Furthermore, the compound can be administered as a component of a microcapsule that can be targeted to specific cell types, such as macrophages, or where the diffusion of the compound or delivery of the compound from the microcapsule is designed for a specific rate or dosage.

320. In the methods described above which include the administration and uptake of exogenous DNA into the cells of a subject (i.e., gene transduction or transfection), delivery of the compositions to cells can be via a variety of mechanisms. As one example, delivery can be via a liposome, using commercially available liposome preparations such as LIPOFECTIN, LIPOFECTAMINE (GIBCO-BRL, Inc., Gaithersburg, MD), SUPERFECT (Qiagen, Inc. Hilden, Germany) and TRANSFECTAM (Promega Biotec, Inc., Madison, WI), as well as other liposomes developed according to procedures standard in the art. In addition, the nucleic acid or vector of this invention can be delivered *in vivo* by electroporation, the technology for which is available from Genetronics, Inc. (San Diego, CA) as well as by means of a SONOPORATION machine (ImaRx Pharmaceutical Corp., Tucson, AZ).

321. The materials may be in solution, suspension (for example, incorporated into microparticles, liposomes, or cells). These may be targeted to a particular cell type via antibodies, receptors, or receptor ligands. The following references are examples of the use of this technology to target specific proteins to tumor tissue (Senter, et al., Bioconjugate Chem., 2:447-451, (1991); Bagshawe, K.D., Br. J. Cancer, 60:275-281, (1989); Bagshawe, et al., Br. J. Cancer, 58:700-703, (1988); Senter, et al., Bioconjugate Chem., 4:3-9, (1993); Battelli, et al., Cancer Immunol. Immunother., 35:421-425, (1992); Pietersz and McKenzie, Immunolog. Reviews, 129:57-80, (1992); and Roffler, et al., Biochem. Pharmacol, 42:2062-2065, (1991)). These techniques can be used for a variety of other specific cell types. Vehicles such as "stealth" and other antibody conjugated liposomes (including lipid mediated drug targeting to colonic carcinoma), receptor mediated targeting of DNA through cell specific ligands, lymphocyte directed tumor targeting, and highly specific therapeutic retroviral targeting of murine glioma cells *in vivo*. The following references are examples of the use of this technology to target specific proteins to tumor tissue (Hughes et al., Cancer Research, 49:6214-6220, (1989); and Litzinger and Huang, Biochimica et Biophysica Acta, 1104:179-187, (1992)). In general, receptors are involved in pathways of endocytosis, either constitutive or ligand induced. These receptors cluster in clathrin-coated pits, enter the cell via clathrin-coated vesicles, pass through an acidified endosome in which the receptors are sorted, and then either recycle to the cell surface, become stored intracellularly, or are degraded in lysosomes. The internalization pathways serve a variety of functions, such as nutrient uptake, removal of

activated proteins, clearance of macromolecules, opportunistic entry of viruses and toxins, dissociation and degradation of ligand, and receptor-level regulation. Many receptors follow more than one intracellular pathway, depending on the cell type, receptor concentration, type of ligand, ligand valency, and ligand concentration. Molecular and cellular mechanisms of receptor-mediated endocytosis has been reviewed (Brown and Greene, DNA and Cell Biology 10:6, 399-409 (1991)).

(3) In vivo/ex vivo

322. As described above, the compositions can be administered in a pharmaceutically acceptable carrier and can be delivered to the subjects cells *in vivo* and/or *ex vivo* by a variety of mechanisms well known in the art (e.g., uptake of naked DNA, liposome fusion, intramuscular injection of DNA via a gene gun, endocytosis and the like).

323. If *ex vivo* methods are employed, cells or tissues can be removed and maintained outside the body according to standard protocols well known in the art. The compositions can be introduced into the cells via any gene transfer mechanism, such as, for example, calcium phosphate mediated gene delivery, electroporation, microinjection or proteoliposomes. The transduced cells can then be infused (e.g., in a pharmaceutically acceptable carrier) or homotopically transplanted back into the subject per standard methods for the cell or tissue type. Standard methods are known for transplantation or infusion of various cells into a subject.

e) Expression systems

324. The nucleic acids that are delivered to cells typically contain expression controlling systems. For example, the inserted genes in viral and retroviral systems usually contain promoters, and/or enhancers to help control the expression of the desired gene product. A promoter is generally a sequence or sequences of DNA that function when in a relatively fixed location in regard to the transcription start site. A promoter contains core elements required for basic interaction of RNA polymerase and transcription factors, and may contain upstream elements and response elements.

(1) Viral Promoters and Enhancers

325. Preferred promoters controlling transcription from vectors in mammalian host cells may be obtained from various sources, for example, the genomes of viruses such as: polyoma, Simian Virus 40 (SV40), adenovirus, retroviruses, hepatitis-B virus and most preferably cytomegalovirus, or from heterologous mammalian promoters, e.g. beta actin promoter. The early and late promoters of the SV40 virus are conveniently obtained as an SV40 restriction fragment which also contains the SV40 viral origin of replication (Fiers et al., Nature, 273: 113 (1978)). The immediate early promoter of the human cytomegalovirus is conveniently obtained as a HindIII E restriction fragment (Greenway, P.J. et al., Gene 18: 355-360 (1982)). Of course, promoters from the host cell or related species also are useful herein.

326. Enhancer generally refers to a sequence of DNA that functions at no fixed distance from the transcription start site and can be either 5' (Laimins, L. et al., Proc. Natl. Acad. Sci. 78: 993 (1981)) or 3' (Lusky, M.L., et al., Mol. Cell Bio. 3: 1108 (1983)) to the transcription unit. Furthermore, enhancers can be within an intron (Banerji, J.L. et al., Cell 33: 729 (1983)) as well as within the coding sequence itself

(Osborne, T.F., et al., Mol. Cell Bio. 4: 1293 (1984)). They are usually between 10 and 300 bp in length, and they function in cis. Enhancers f unction to increase transcription from nearby promoters. Enhancers also often contain response elements that mediate the regulation of transcription. Promoters can also contain response elements that mediate the regulation of transcription. Enhancers often determine the regulation of expression of a gene. While many enhancer sequences are now known from mammalian genes (globin, elastase, albumin, -fetoprotein and insulin), typically one will use an enhancer from a eukaryotic cell virus for general expression. Examples are the SV40 enhancer on the late side of the replication origin (bp 100-270), the cytomegalovirus early promoter enhancer, the polyoma enhancer on the late side of the replication origin, and adenovirus enhancers.

327. The promotor and/or enhancer may be specifically activated either by light or specific chemical events which trigger their function. Systems can be regulated by reagents such as tetracycline and dexamethasone. There are also ways to enhance viral vector gene expression by exposure to irradiation, such as gamma irradiation, or alkylating chemotherapy drugs.

328. In certain embodiments the promoter and/or enhancer region can act as a constitutive promoter and/or enhancer to maximize expression of the region of the transcription unit to be transcribed. In certain constructs the promoter and/or enhancer region be active in all eukaryotic cell types, even if it is only expressed in a particular type of cell at a particular time. A preferred promoter of this type is the CMV promoter (650 bases). Other preferred promoters are SV40 promoters, cytomegalovirus (full length promoter), and retroviral vector LTF.

329. It has been shown that all specific regulatory elements can be cloned and used to construct expression vectors that are selectively expressed in specific cell types such as melanoma cells. The glial fibrillary acetic protein (GFAP) promoter has been used to selectively express genes in cells of glial origin.

330. Expression vectors used in eukaryotic host cells (yeast, fungi, insect, plant, animal, human or nucleated cells) may also contain sequences necessary for the termination of transcription which may affect mRNA expression. These regions are transcribed as polyadenylated segments in the untranslated portion of the mRNA encoding tissue factor protein. The 3' untranslated regions also include transcription termination sites. It is preferred that the transcription unit also contain a polyadenylation region. One benefit of this region is that it increases the likelihood that the transcribed unit will be processed and transported like mRNA. The identification and use of polyadenylation signals in expression constructs is well established. It is preferred that homologous polyadenylation signals be used in the transgene constructs. In certain transcription units, the polyadenylation region is derived from the SV40 early polyadenylation signal and consists of about 400 bases. It is also preferred that the transcribed units contain other standard sequences alone or in combination with the above sequences improve expression from, or stability of, the construct.

(2) Markers

331. The vectors can include nucleic acid sequence encoding a marker product. This marker product is used to determine if the gene has been delivered to the cell and once delivered is being expressed. Preferred marker genes are the *E. Coli* lacZ gene, which encodes ß-galactosidase, and green fluorescent protein.

332. In some embodiments the marker may be a selectable marker. Examples of suitable selectable markers for mammalian cells are dihydrofolate reductase (DHFR), thymidine kinase, neomycin, neomycin analog G418, hydromycin, and puromycin. When such selectable markers are successfully transferred into a mammalian host cell, the transformed mammalian host cell can survive if placed under selective pressure. There are two widely used distinct categories of selective regimes. The first category is based on a cell's metabolism and the use of a mutant cell line which lacks the ability to grow independent of a supplemented media. Two examples are: CHO DHFR- cells and mouse LTK- cells. These cells lack the ability to grow without the addition of such nutrients as thymidine or hypoxanthine. Because these cells lack certain genes necessary for a complete nucleotide synthesis pathway, they cannot survive unless the missing nucleotides are provided in a supplemented media. An alternative to supplementing the media is to introduce an intact DHFR or TK gene into cells lacking the respective genes, thus altering their growth requirements. Individual cells which were not transformed with the DHFR or TK gene will not be capable of survival in non-supplemented media.

333. The second category is dominant selection which refers to a selection scheme used in any cell type and does not require the use of a mutant cell line. These schemes typically use a drug to arrest growth of a host cell. Those cells which have a novel gene would express a protein conveying drug resistance and would survive the selection. Examples of such dominant selection use the drugs neomycin, (Southern P. and Berg, P., J. Molec. Appl. Genet. 1: 327 (1982)), mycophenolic acid, (Mulligan, R.C. and Berg, P. Science 209: 1422 (1980)) or hygromycin, (Sugden, B. et al., Mol. Cell. Biol. 5: 410-413 (1985)). The three examples employ bacterial genes under eukaryotic control to convey resistance to the appropriate drug G418 or neomycin (geneticin), xgpt (mycophenolic acid) or hygromycin, respectively. Others include the neomycin analog G418 and puramycin.

f) Peptides

(1) Protein variants

334. As discussed herein there are numerous variants of the TSG101, Vps4A, or Vps4B proteins and GAG protein that are known and herein contemplated. In addition, to the known functional TSG101, Vps4A, or Vps4B strains and GAG strain there are derivatives and variants of the TSG101, Vps4A, or Vps4B and GAG proteins which also function in the disclosed methods and compositions. Protein variants and derivatives are well understood to those of skill in the art and in can involve amino acid sequence modifications. For example, amino acid sequence modifications typically fall into one or more of three classes: substitutional, insertional or deletional variants. Insertions include amino and/or carboxyl terminal fusions as well as intrasequence insertions of single or multiple amino acid residues. Insertions ordinarily will be smaller insertions than those of amino or carboxyl terminal fusions, for

example, on the order of one to four residues. Immunogenic fusion protein derivatives, such as those described in the examples, are made by fusing a polypeptide sufficiently large to confer immunogenicity to the target sequence by cross-linking in vitro or by recombinant cell culture transformed with DNA encoding the fusion. Deletions are characterized by the removal of one or more amino acid residues from the protein sequence. Typically, no more than about from 2 to 6 residues are deleted at any one site within the protein molecule. These variants ordinarily are prepared by site specific mutagenesis of nucleotides in the DNA encoding the protein, thereby producing DNA encoding the variant, and thereafter expressing the DNA in recombinant cell culture. Techniques for making substitution mutations at predetermined sites in DNA having a known sequence are well known, for example M13 primer mutagenesis and PCR mutagenesis. Amino acid substitutions are typically of single residues, but can occur at a number of different locations at once; insertions usually will be on the order of about from 1 to 10 amino acid residues; and deletions will range about from 1 to 30 residues. Deletions or insertions preferably are made in adjacent pairs, i.e. a deletion of 2 residues or insertion of 2 residues. Substitutions, deletions, insertions or any combination thereof may be combined to arrive at a final construct. The mutations must not place the sequence out of reading frame and preferably will not create complementary regions that could produce secondary mRNA structure. Substitutional variants are those in which at least one residue has been removed and a different residue inserted in its place. Such substitutions generally are made in accordance with the following Tables 1 and 2 and are referred to as conservative substitutions.

335. TABLE 1:Amino Acid Abbreviations

Amino Acid	Abbreviations
Alanine	AlaA
Allosoleucine	AIle
Arginine	ArgR
Asparagines	AsnN
aspartic acid	AspD
Cysteine	CysC
glutamic acid	GluE
Glutamine	GlnQ
Glycine	GlyG
Histidine	HisH
Isolelucine	IleI
Leucine	LeuL
Lysine	LysK
Phenylalanine	PheF
Proline	ProP
pyroglutamic acidp	Glu
Serine	SerS
Threonine	ThrT
Tyrosine	TyrY
Tryptophan	TrpW
Valine	ValV

Table 2.

Original Residue Exemplary Conservative Substitutions, others are known in the art.

Ala	ser
Arg	lys, gln

Asn	gln; his
Asp	glu
Cys	ser
Gln	asn, lys
Glu	asp
Gly	pro
His	asn;gln
Ile	leu; val
Leu	ile; val
Lys	arg; gln;
Met	Leu; ile
Phe	met; leu; tyr
Ser	thr
Thr	ser
Trp	tyr
Tyr	trp; phe
Val	ile; leu

336. Substantial changes in function or immunological identity are made by selecting substitutions that are less conservative than those in Table 2, i.e., selecting residues that differ more significantly in their effect on maintaining (a) the structure of the polypeptide backbone in the area of the substitution, for example as a sheet or helical conformation, (b) the charge or hydrophobicity of the molecule at the target site or (c) the bulk of the side chain. The substitutions which in general are expected to produce the greatest changes in the protein properties will be those in which (a) a hydrophilic residue, e.g. seryl or threonyl, is substituted for (or by) a hydrophobic residue, e.g. leucyl, isoleucyl, phenylalanyl, valyl or alanyl; (b) a cysteine or proline is substituted for (or by) any other residue; (c) a residue having an electropositive side chain, e.g., lysyl, arginyl, or histidyl, is substituted for (or by) an electronegative residue, e.g., glutamyl or aspartyl; or (d) a residue having a bulky side chain, e.g., phenylalanine, is substituted for (or by) one not having a side chain, e.g., glycine, in this case, (e) by increasing the number of sites for sulfation and/or glycosylation.

337. For example, the replacement of one amino acid residue with another that is biologically and/or chemically similar is known to those skilled in the art as a conservative substitution. For example, a conservative substitution would be replacing one hydrophobic residue for another, or one polar residue for another. The substitutions include combinations such as, for example, Gly, Ala; Val, Ile, Leu; Asp, Glu; Asn, Gln; Ser, Thr; Lys, Arg; and Phe, Tyr. Such conservatively substituted variations of each explicitly disclosed sequence are included within the mosaic polypeptides provided herein.

338. Substitutional or deletional mutagenesis can be employed to insert sites for N-glycosylation (Asn-X-Thr/Ser) or O-glycosylation (Ser or Thr). Deletions of cysteine or other labile residues also may be desirable. Deletions or substitutions of potential proteolysis sites, e.g. Arg, is accomplished for example by deleting one of the basic residues or substituting one by glutaminyl or histidyl residues.

339. Certain post-translational derivatizations are the result of the action of recombinant host cells on the expressed polypeptide. Glutaminyl and asparaginyl residues are frequently post-translationally deamidated to the corresponding glutamyl and asparyl residues. Alternatively, these residues are deamidated under mildly acidic conditions. Other post-translational modifications include

hydroxylation of proline and lysine, phosphorylation of hydroxyl groups of seryl or threonyl residues, methylation of the o-amino groups of lysine, arginine, and histidine side chains (T.E. Creighton, Proteins: Structure and Molecular Properties, W. H. Freeman & Co., San Francisco pp 79-86 [1983]), acetylation of the N-terminal amine and, in some instances, amidation of the C-terminal carboxyl.

340. It is understood that one way to define the variants and derivatives of the disclosed proteins herein is through defining the variants and derivatives in terms of homology/identity to specific known sequences. For example, SEQ ID NO:1 sets forth a particular sequence of TSG101 and SEQ ID NO:5 sets forth a particular sequence of a retroviral GAG protein. Specifically disclosed are variants of these and other proteins herein disclosed which have at least, 70% or 75% or 80% or 85% or 90% or 95% homology to the stated sequence. Those of skill in the art readily understand how to determine the homology of two proteins. For example, the homology can be calculated after aligning the two sequences so that the homology is at its highest level.

341. Another way of calculating homology can be performed by published algorithms. Optimal alignment of sequences for comparison may be conducted by the local homology algorithm of Smith and Waterman Adv. Appl. Math. 2: 482 (1981), by the homology alignment algorithm of Needleman and Wunsch, J. MoL Biol. 48: 443 (1970), by the search for similarity method of Pearson and Lipman, Proc. Natl. Acad. Sci. U.S.A. 85: 2444 (1988), by computerized implementations of these algorithms (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, WI), or by inspection.

342. The same types of homology can be obtained for nucleic acids by for example the algorithms disclosed in Zuker, M. *Science* 244:48-52, 1989, Jaeger et al. *Proc. Natl. Acad. Sci. USA* 86:7706-7710, 1989, Jaeger et al. *Methods Enzymol.* 183:281-306, 1989 which are herein incorporated by reference for at least material related to nucleic acid alignment.

343. It is understood that the description of conservative mutations and homology can be combined together in any combination, such as embodiments that have at least 70% homology to a particular sequence wherein the variants are conservative mutations.

344. As this specification discusses various proteins and protein sequences it is understood that the nucleic acids that can encode those protein sequences are also disclosed. This would include all degenerate sequences related to a specific protein sequence, i.e. all nucleic acids having a sequence that encodes one particular protein sequence as well as all nucleic acids, including degenerate nucleic acids, encoding the disclosed variants and derivatives of the protein sequences. Thus, while each particular nucleic acid sequence may not be written out herein, it is understood that each and every sequence is in fact disclosed and described herein through the disclosed protein sequence. It is also understood that while no amino acid sequence indicates what particular DNA sequence encodes that protein within an organism, where particular variants of a disclosed protein are disclosed herein, the known nucleic acid sequence that encodes that protein in the particular organism from which that protein arises is also known and herein disclosed and described.

g) Pharmaceutical carriers/Delivery of pharamceutical products

345. As described herein, the compositions can also be administered *in vivo* in a pharmaceutically acceptable carrier. By "pharmaceutically acceptable" is meant a material that is not biologically or otherwise undesirable, i.e., the material may be administered to a subject, along with the nucleic acid or vector, without causing any undesirable biological effects or interacting in a deleterious manner with any of the other components of the pharmaceutical composition in which it is contained. The carrier would naturally be selected to minimize any degradation of the active ingredient and to minimize any adverse side effects in the subject, as would be well known to one of skill in the art.

346. The compositions may be administered orally, parenterally (e.g., intravenously), by intramuscular injection, by intraperitoneal injection, transdermally, extracorporeally, topically or the like, including topical intranasal administration or administration by inhalant. As used herein, "topical intranasal administration" means delivery of the compositions into the nose and nasal passages through one or both of the nares and can comprise delivery by a spraying mechanism or droplet mechanism, or through aerosolization of the composition. Administration of the compositions by inhalant can be through the nose or mouth via delivery by a spraying or droplet mechanism. Delivery can also be directly to any area of the respiratory system (e.g., lungs) via intubation. The exact amount of the compositions required will vary from subject to subject, depending on the species, age, weight and general condition of the subject, the severity of the allergic disorder being treated, the particular nucleic acid or vector used, its mode of administration and the like. Thus, it is not possible to specify an exact amount for every composition. However, an appropriate amount can be determined by one of ordinary skill in the art using only routine experimentation given the teachings herein.

347. Parenteral administration of the composition, if used, is generally characterized by injection. Injectables can be prepared in conventional forms, either as liquid solutions or suspensions, solid forms suitable for solution of suspension in liquid prior to injection, or as emulsions. A more recently revised approach for parenteral administration involves use of a slow release or sustained release system such that a constant dosage is maintained. See, e.g., U.S. Patent No. 3,610,795, which is incorporated by reference herein.

348. The materials may be in solution, suspension (for example, incorporated into microparticles, liposomes, or cells). These may be targeted to a particular cell type via antibodies, receptors, or receptor ligands. The following references are examples of the use of this technology to target specific proteins to tumor tissue (Senter, et al., Bioconjugate Chem., 2:447-451, (1991); Bagshawe, K.D., Br. J. Cancer, 60:275-281, (1989); Bagshawe, et al., Br. J. Cancer, 58:700-703, (1988); Senter, et al., Bioconjugate Chem., 4:3-9, (1993); Battelli, et al., Cancer Immunol. Immunother., 35:421-425, (1992); Pietersz and McKenzie, Immunolog. Reviews, 129:57-80, (1992); and Roffler, et al., Biochem. Pharmacol, 42:2062-2065, (1991)). Vehicles such as "stealth" and other antibody conjugated liposomes (including lipid mediated drug targeting to colonic carcinoma), receptor mediated targeting of DNA through cell specific ligands, lymphocyte directed tumor targeting, and highly specific therapeutic retroviral targeting of

murine glioma cells *in vivo*. The following references are examples of the use of this technology to target specific proteins to tumor tissue (Hughes et al., Cancer Research, 49:6214-6220, (1989); and Litzinger and Huang, Biochimica et Biophysica Acta, 1104:179-187, (1992)). In general, receptors are involved in pathways of endocytosis, either constitutive or ligand induced. These receptors cluster in clathrin-coated pits, enter the cell via clathrin-coated vesicles, pass through an acidified endosome in which the receptors are sorted, and then either recycle to the cell surface, become stored intracellularly, or are degraded in lysosomes. The internalization pathways serve a variety of functions, such as nutrient uptake, removal of activated proteins, clearance of macromolecules, opportunistic entry of viruses and toxins, dissociation and degradation of ligand, and receptor-level regulation. Many receptors follow more than one intracellular pathway, depending on the cell type, receptor concentration, type of ligand, ligand valency, and ligand concentration. Molecular and cellular mechanisms of receptor-mediated endocytosis has been reviewed (Brown and Greene, DNA and Cell Biology 10:6, 399-409 (1991)).

(1) Pharmaceutically Acceptable Carriers

349. The compositions, including antibodies or other compositions which interact with any of the disclosed compositions, such as TSG101, Vps4A, or Vps4B such that TSG101, Vps4A, or Vps4B cannot interact with the retroviral GAG and/or inhibit retroviral budding, can be used therapeutically in combination with a pharmaceutically acceptable carrier.

350. Suitable carriers and their formulations are described in *Remington: The Science and Practice of Pharmacy* (19th ed.) ed. A.R. Gennaro, Mack Publishing Company, Easton, PA 1995. Typically, an appropriate amount of a pharmaceutically-acceptable salt is used in the formulation to render the formulation isotonic. Examples of the pharmaceutically-acceptable carrier include, but are not limited to, saline, Ringer's solution and dextrose solution. The pH of the solution is preferably from about 5 to about 8, and more preferably from about 7 to about 7.5. Further carriers include sustained release preparations such as semipermeable matrices of solid hydrophobic polymers containing the antibody, which matrices are in the form of shaped articles, e.g., films, liposomes or microparticles. It will be apparent to those persons skilled in the art that certain carriers may be more preferable depending upon, for instance, the route of administration and concentration of composition being administered.

351. Pharmaceutical carriers are known to those skilled in the art. These most typically would be standard carriers for administration of drugs to humans, including solutions such as sterile water, saline, and buffered solutions at physiological pH. The compositions can be administered intramuscularly or subcutaneously. Other compounds will be administered according to standard procedures used by those skilled in the art.

352. Pharmaceutical compositions may include carriers, thickeners, diluents, buffers, preservatives, surface active agents and the like in addition to the molecule of choice. Pharmaceutical compositions may also include one or more active ingredients such as antimicrobial agents, antiinflammatory agents, anesthetics, and the like.

353. The pharmaceutical composition may be administered in a number of ways depending on whether local or systemic treatment is desired, and on the area to be treated. Administration may be topically (including ophthalmically, vaginally, rectally, intranasally), orally, by inhalation, or parenterally, for example by intravenous drip, subcutaneous, intraperitoneal or intramuscular injection. The disclosed antibodies can be administered intravenously, intraperitoneally, intramuscularly, subcutaneously, intracavity, or transdermally.

354. Preparations for parenteral administration include sterile aqueous or non-aqueous solutions, suspensions, and emulsions. Examples of non-aqueous solvents are propylene glycol, polyethylene glycol, vegetable oils such as olive oil, and injectable organic esters such as ethyl oleate. Aqueous carriers include water, alcoholic/aqueous solutions, emulsions or suspensions, including saline and buffered media. Parenteral vehicles include sodium chloride solution, Ringer's dextrose, dextrose and sodium chloride, lactated Ringer's, or fixed oils. Intravenous vehicles include fluid and nutrient replenishers, electrolyte replenishers (such as those based on Ringer's dextrose), and the like. Preservatives and other additives may also be present such as, for example, antimicrobials, anti-oxidants, chelating agents, and inert gases and the like.

355. Formulations for topical administration may include ointments, lotions, creams, gels, drops, suppositories, sprays, liquids and powders. Conventional pharmaceutical carriers, aqueous, powder or oily bases, thickeners and the like may be necessary or desirable.

356. Compositions for oral administration include powders or granules, suspensions or solutions in water or non-aqueous media, capsules, sachets, or tablets. Thickeners, flavorings, diluents, emulsifiers, dispersing aids or binders may be desirable.

357. Some of the compositions may potentially be administered as a pharmaceutically acceptable acid- or base- addition salt, formed by reaction with inorganic acids such as hydrochloric acid, hydrobromic acid, perchloric acid, nitric acid, thiocyanic acid, sulfuric acid, and phosphoric acid, and organic acids such as formic acid, acetic acid, propionic acid, glycolic acid, lactic acid, pyruvic acid, oxalic acid, malonic acid, succinic acid, maleic acid, and fumaric acid, or by reaction with an inorganic base such as sodium hydroxide, ammonium hydroxide, potassium hydroxide, and organic bases such as mono-, di-, trialkyl and aryl amines and substituted ethanolamines.

(2) Therapeutic Uses

358. Effective dosages and schedules for administering the compositions may be determined empirically, and making such determinations is within the skill in the art. The dosage ranges for the administration of the compositions are those large enough to produce the desired effect in which the symptoms disorder are effected. The dosage should not be so large as to cause adverse side effects, such as unwanted cross-reactions, anaphylactic reactions, and the like. Generally, the dosage will vary with the age, condition, sex and extent of the disease in the patient, route of administration, or whether other drugs are included in the regimen, and can be determined by one of skill in the art. The dosage can be adjusted by the individual physician in the event of any counterindications. Dosage can vary, and can be

administered in one or more dose administrations daily, for one or several days. Guidance can be found in the literature for appropriate dosages for given classes of pharmaceutical products. For example, guidance in selecting appropriate doses for antibodies can be found in the literature on therapeutic uses of antibodies, e.g., Handbook of Monoclonal Antibodies, Ferrone et al., eds., Noges Publications, Park Ridge, N.J., (1985) ch. 22 and pp. 303-357; Smith et al., Antibodies in Human Diagnosis and Therapy, Haber et al., eds., Raven Press, New York (1977) pp. 365-389. A typical daily dosage of the antibody used alone might range from about 1 μg/kg to up to 100 mg/kg of body weight or more per day, depending on the factors mentioned above.

359. Following administration of a disclosed composition, such as an antibody or molecule that inhibits, for example, TSG101, Vps4A, or Vps4B from interacting with GAG, for treating, inhibiting, or preventing an HIV infection, the efficacy of the therapeutic composition can be assessed in various ways well known to the skilled practitioner. For instance, one of ordinary skill in the art will understand that the compositions disclosed herein are efficacious in treating or inhibiting an HIV infection in a subject by observing that the composition reduces viral load or prevents a further increase in viral load. Viral loads can be measured by methods that are known in the art, for example, using polymerase chain reaction assays to detect the presence of HIV nucleic acid or antibody assays to detect the presence of HIV protein in a sample (e.g., but not limited to, blood) from a subject or patient, or by measuring the level of circulating anti-HIV antibody levels in the patient. Efficacy of the administration of the disclosed composition may also be determined by measuring the number of $CD4^+$ T cells in the HIV-infected subject. An antibody treatment that inhibits an initial or further decrease in $CD4^+$ T cells in an HIV-positive subject or patient, or that result in an increase in the number of $CD4^+$ T cells in the HIV-positive subject, is an efficacious antibody treatment.

360. The compositions that inhibit, for example, TSG101-PTAP interactions or TSG101, Vps4A or Vps4B interactions disclosed herein may be administered prophylactically to patients or subjects who are at risk for being exposed to HIV or who have been newly exposed to HIV. In subjects who have been newly exposed to HIV but who have not yet displayed the presence of the virus (as measured by PCR or other assays for detecting the virus) in blood or other body fluid, efficacious treatment with an antibody of the invention partially or completely inhibits the appearance of the virus in the blood or other body fluid.

361. Other molecules that interact with, for example, TSG101, Vps4A, or Vps4B to inhibit TSG101, Vps4A or Vps4B interactions which do not have a specific pharmaceutical function, but which may be used for tracking changes within cells or for the delivery of diagnostic tools for example can be delivered in ways similar to those described for the pharmaceutical products.

362. The disclosed compositions and methods can also be used for example as tools to isolate and test new drug candidates for a variety of PTAP motif containing retroviral related diseases. They can also be used for the continued isolation and study, for example, the cell cycle.

7. **Chips and micro arrays**

363. Disclosed are chips where at least one address is the sequences or part of the sequences set forth in any of the nucleic acid sequences disclosed herein. Also disclosed are chips where at least one address is the sequences or portion of sequences set forth in any of the peptide sequences disclosed herein.

364. Also disclosed are chips where at least one address is a variant of the sequences or part of the sequences set forth in any of the nucleic acid sequences disclosed herein. Also disclosed are chips where at least one address is a variant of the sequences or portion of sequences set forth in any of the peptide sequences disclosed herein.

8. **Kits**

365. Disclosed herein are kits that are drawn to reagents that can be used in practicing the methods disclosed herein. The kits can include any reagent or combination of reagent discussed herein or that would be understood to be required or beneficial in the practice of the disclosed methods. For example, the kits could include primers to perform the amplification reactions discussed in certain embodiments of the methods, as well as the buffers and enzymes required to use the primers as intended.

D. Methods of making the compositions

366. The compositions disclosed herein and the compositions necessary to perform the disclosed methods can be made using any method known to those of skill in the art for that particular reagent or compound unless otherwise specifically noted.

1. **Nucleic acid synthesis**

367. For example, the nucleic acids, such as, the oligonucleotides to be used as primers can be made using standard chemical synthesis methods or can be produced using enzymatic methods or any other known method. Such methods can range from standard enzymatic digestion followed by nucleotide fragment isolation (see for example, Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*, 2nd Edition (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1989) Chapters 5, 6) to purely synthetic methods, for example, by the cyanoethyl phosphoramidite method using a Milligen or Beckman System 1Plus DNA synthesizer (for example, Model 8700 automated synthesizer of Milligen-Biosearch, Burlington, MA or ABI Model 380B). Synthetic methods useful for making oligonucleotides are also described by Ikuta *et al.*, *Ann. Rev. Biochem.* **53**:323-356 (1984), (phosphotriester and phosphite-triester methods), and Narang *et al.*, *Methods Enzymol.*, 65:610-620 (1980), (phosphotriester method). Protein nucleic acid molecules can be made using known methods such as those described by Nielsen *et al.*, *Bioconjug. Chem.* **5**:3-7 (1994).

2. **Peptide synthesis**

368. One method of producing the disclosed proteins, such as SEQ ID NO:23, is to link two or more peptides or polypeptides together by protein chemistry techniques. For example, peptides or polypeptides can be chemically synthesized using currently available laboratory equipment using either Fmoc (9-fluorenylmethyloxycarbonyl) or Boc (*tert*-butyloxycarbonoyl) chemistry. (Applied Biosystems, Inc., Foster City, CA). One skilled in the art can readily appreciate that a peptide or

polypeptide corresponding to the disclosed proteins, for example, can be synthesized by standard chemical reactions. For example, a peptide or polypeptide can be synthesized and not cleaved from its synthesis resin whereas the other fragment of a peptide or protein can be synthesized and subsequently cleaved from the resin, thereby exposing a terminal group which is functionally blocked on the other fragment. By peptide condensation reactions, these two fragments can be covalently joined via a peptide bond at their carboxyl and amino termini, respectively, to form an antibody, or fragment thereof. (Grant GA (1992) Synthetic Peptides: A User Guide. W.H. Freeman and Co., N.Y. (1992); Bodansky M and Trost B., Ed. (1993) Principles of Peptide Synthesis. Springer-Verlag Inc., NY (which is herein incorporated by reference at least for material related to peptide synthesis). Alternatively, the peptide or polypeptide is independently synthesized *in vivo* as described herein. Once isolated, these independent peptides or polypeptides may be linked to form a peptide or fragment thereof via similar peptide condensation reactions.

369. For example, enzymatic ligation of cloned or synthetic peptide segments allow relatively short peptide fragments to be joined to produce larger peptide fragments, polypeptides or whole protein domains (Abrahmsen L et al., Biochemistry, 30:4151 (1991)). Alternatively, native chemical ligation of synthetic peptides can be utilized to synthetically construct large peptides or polypeptides from shorter peptide fragments. This method consists of a two step chemical reaction (Dawson et al. Synthesis of Proteins by Native Chemical Ligation. Science, 266:776-779 (1994)). The first step is the chemoselective reaction of an unprotected synthetic peptide--thioester with another unprotected peptide segment containing an amino-terminal Cys residue to give a thioester-linked intermediate as the initial covalent product. Without a change in the reaction conditions, this intermediate undergoes spontaneous, rapid intramolecular reaction to form a native peptide bond at the ligation site (Baggiolini M et al. (1992) FEBS Lett. 307:97-101; Clark-Lewis I et al., J.Biol.Chem., 269:16075 (1994); Clark-Lewis I et al., Biochemistry, 30:3128 (1991); Rajarathnam K et al., Biochemistry 33:6623-30 (1994)).

370. Alternatively, unprotected peptide segments are chemically linked where the bond formed between the peptide segments as a result of the chemical ligation is an unnatural (non-peptide) bond (Schnolzer, M et al. Science, 256:221 (1992)). This technique has been used to synthesize analogs of protein domains as well as large amounts of relatively pure proteins with full biological activity (deLisle Milton RC et al., Techniques in Protein Chemistry IV. Academic Press, New York, pp. 257-267 (1992)).

3. Process for making the compositions

371. Disclosed are processes for making the compositions as well as making the intermediates leading to the compositions. There are a variety of methods that can be used for making these compositions, such as synthetic chemical methods and standard molecular biology methods. It is understood that the methods of making these and the other disclosed compositions are specifically disclosed.

E. Methods of using the compositions

1. Methods of using the compositions as research tools

372. The disclosed compositions can be used in a variety of ways as research tools. For example, the disclosed coordinates and storage media containing them, can be used to isolate molecules that interact with the TSG101, Vps4A, or Vps4B or GAG molecule.

373. The compositions can be used for example as targets in combinatorial chemistry protocols or other screening protocols to isolate molecules that possess desired functional properties related to binding TSG101, Vps4A, or Vps4B, such that TSG101, Vps4A, or Vps4B cannot interact with retroviral GAG proteins and/or inhibit retroviral budding.

F. Examples

374. The following examples are put forth so as to provide those of ordinary skill in the art with a complete disclosure and description of how the compounds, compositions, articles, devices and/or methods claimed herein are made and evaluated, and are intended to be purely exemplary of the invention and are not intended to limit the scope of what the inventors regard as their invention. Efforts have been made to ensure accuracy with respect to numbers (e.g., amounts, temperature, etc.), but some errors and deviations should be accounted for. Unless indicated otherwise, parts are parts by weight, temperature is in °C or is at ambient temperature, and pressure is at or near atmospheric.

1. Example 1 Tsg101, Vps4, and HIV-1 Budding

a) Results

(1) Human Tsg101 Binds Specifically to the PTAP Motif of HIV-1 p6

375. Disclosed is a strategy to screen for proteins that could bind to the PTAP late domain within HIV-1 p6 and obtained similar results. Briefly, HIV-$1_{NYU/BRS}$ p6 was used as bait to screen a human spleen cDNA prey library for potential binding partners. Genes encoding nearly full length Tsg101 were isolated twice, and were the only genes detected and confirmed in these screens. In subsequent experiments, it was shown that full length Tsg101 bound wild type p6 in directed two-hybrid liquid culture assays, resulting in high levels of β-galactosidase activity (>300-fold over background; not shown). Three different p6 point mutants (P7L, A9R and P10L) were used to test whether the Tsg101 binding interaction required the PTAP motif within HIV-1 p6, and all three reduced β-galactosidase activity to background levels. Each of these point mutations also arrests HIV-1 budding at a late stage (Huang et al., 1995). Hence, Tsg101 was judged an attractive candidate for the cellular factor that binds the PTAP late domain of HIV-1 p6 and facilitates viral budding.

376. Directed yeast two-hybrid experiments were further used to map the primary p6 binding site to the N-terminal half of Tsg101 (residues 1-207, data not shown). Comparative sequence analyses have suggested that this region of Tsg101 contains a domain that belongs to the ubiquitin enzyme 2 variant (UEV) protein family (Koonin and Abagyan, 1997; Ponting et al., 1997; VanDemark et al., 2001). UEV proteins are homologous to the E2 class of ubiquitin conjugating enzymes, but lack the active site cysteine residue required to make a transient thioester bond during ubiquitin transfer, and are presumably

not enzymatically active. A construct spanning Tsg101 residues 1-145 was used to characterize the interaction between HIV-1_{NL4-3} p6 and the Tsg101 UEV domain *in vitro*. Recombinant Tsg101 UEV was expressed in *E. coli*, purified to homogeneity, and tested for binding to immobilized GST-p6 fusion proteins using a surface plasmon resonance biosensor (Fig. 1). Tsg101 UEV exhibited concentration-dependent p6 binding, and equilibrium responses fit a simple 1:1 binding model with an equilibrium dissociation constant ($K_d^{20°C}$) of 27±5 μM. As summarized in Fig. 1A, a series of constructs spanning p6 residues 1-13, 1-27, 1-33, and 1-52 (full length p6) all bound Tsg101 UEV with similar affinities. Thus, the primary Tsg101 binding site is located within the first 13 residues of p6.

377. Site-directed mutagenesis of p6 was then used to map the Tsg101 binding site more precisely. Residues 4-13 within full length p6 were substituted individually with alanine (Fig. 1E, dark bars) or with residues known to block virus release (light bars). Mutations in the four central PTAP residues significantly reduced Tsg101 binding (>2 kcal/mol loss of binding free energy). Mutations at three flanking sites (E6A, P11A and E13A) moderately reduced the binding affinity (0.3-0.5 kcal/mol), indicating that these residues also contributed to Tsg101 UEV binding. Alanine substitutions at three other sites (R4A, P5A, and E12A) had no significant effect on Tsg101 UEV binding. These experiments demonstrate that the Tsg101 UEV domain binds directly and specifically to HIV-1 p6 and that the p6 PTAP motif is the energetically dominant binding epitope.

(2) Ubiquitin Modification of HIV-1 p6 Enhances Tsg101 Binding

378. The enzymes that conjugate Ub onto HIV-1 Gag are not yet known, so models for p6-Ub conjugates were created by fusing the Ub protein in frame to the C-termini of full length p6, $p6_{33}$, and $p6_{27}$ constructs (Fig. 1A). Analogous approaches have been employed by others to rescue Ub-dependent defects in endocytosis (Shih et al., 2000) and RSV budding (Patnaik et al., 2000). In all three p6-Ub constructs, the presence of ubiquitin *increased* the affinity of Tsg101 UEV binding approximately 10-fold (ave. K_d=2.3 μM; Figs. 1A and D). Tsg101 UEV even bound weakly to Ub alone (Fig. 1D, est. K_d=510±35 μM). These experiments demonstrate that Tsg101 UEV, though lacking enzymatic activity, has retained Ub binding activity and can bind cooperatively to ubiquitin and to the HIV-1 p6 late domain. This is consistent with Tsg101 binding ubiquitinated Gag molecules even more tightly *in vivo* if the UEV domain senses the native Ub isopeptide geometry or if oligomerization of full length Tsg101 increases its avidity for the assembling Gag lattice.

(3) Tsg101 is Required for Efficient Release of HIV-1 from 293T Cells

379. It was important to develop systems to test the requirement for cellular Tsg101 in HIV-1 budding. The HIV-1 p6 late domain can mediate virus release from human embryonic kidney 293T cells (Yuan et al., 2000), and this is confirmed herein that a series of mutations in the p6 PTAP sequence that block HIV-1 release and replication in other cell lines also blocked release of infectious HIV-1 particles in 293T cells. Tsg101was depleted from 293T cells using small interfering RNAs (siRNA) (Elbashir et al., 2001). The system was optimized until it became possible to deplete endogenous Tsg101 to nearly undetectable levels (Fig. 2A, lanes 2,8,10). Briefly, 293T cells were transfected twice at 24 h intervals

with a small interfering RNA duplex homologous to nucleotides 413-433 of the Tsg101 coding sequence (denoted siRNA). Tsg101 levels were analyzed after an additional 24 h by Western blotting of cytoplasmic extracts from the bulk culture. Under these conditions, traces of Tsg101 protein could only be detected in highly overexposed Western blots, whereas levels of control proteins were unaffected. Nuclear and membrane fractions were also depleted of Tsg101. A heterologous control RNA duplex of inverted sequence (denoted $siRNA_{INV}$) did not affect cellular Tsg101 levels, demonstrating the specificity of siRNA. Although others have reported that altering Tsg101 levels can affect cell proliferation (Zhong et al., 1998), we observed only a very slight growth reduction in Tsg101-depleted cells over the 72h time course.

380. Tsg101 protein was reintroduced into Tsg101-depleted cells by co-transfecting a plasmid encoding FLAG-tagged Tsg101 protein with 7 silent mutations at the target site that rendered the gene resistant to RNA interference (denoted Tsg*-FLAG). As shown in Figs. 2A and B, this "resistant" Tsg*-FLAG protein was not depleted by siRNA (lanes 4,6). In contrast, endogenous Tsg101 and exogenous Tsg-FLAG expressed from wt genes were always efficiently depleted by siRNA treatment.

381. To test whether Tsg101 was required for HIV-1 budding, Tsg101-depleted cells were co-transfected with a proviral HIV-1 R9 expression construct during the second siRNA transfection. Depletion of Tsg101 very significantly reduced the release of virion-associated MA and CA proteins as analyzed in Western blots, and also reduced viral infectivity in single cycle MAGIC infectivity assays (Figs. 2C and 2E, lanes 2,8,10). Viral titers were reduced 10- to 50-fold in multiple repetitions of this experiment. Both virus release and infectivity were restored to normal levels when the Tsg*-FLAG protein was reintroduced (lanes 4,6), formally demonstrating that Tsg101 is required for the release of infectious HIV-1 from 293T cells. Reductions in virus release in the Tsg101-depleted cells did not reflect general defects in HIV-1 Gag protein synthesis, stability, or processing (Fig. 2D). However, Tsg101 depletion reproducibly increased accumulation of the CA-SP1 Gag processing intermediate in the cytoplasm, consistent with a defect in the final stages of particle assembly (Gottlinger et al., 1991).

382. The effect of Tsg101 depletion on the release of murine leukemia virus (MLV), which uses the alternative "PY" late domain pathway (Yuan et al., 1999) was analyzed. Mutating the MLV Gag p12 PPPY motif (to AAAA) blocked particle release and infectivity (Fig. 3). However, depletion of Tsg101 had only minor, if any, effect on MLV release and infectivity. These experiments demonstrate that Tsg101-depleted 293T cells can still support replication of retroviruses other than HIV-1, provided they are not released via the PTAP-dependent pathway.

(4) A Functional Vps Pathway Is Required for HIV-1 Release from 293T Cells

383. Tsg101 functions in the vacuolar protein sorting (Vps) pathway, in which membrane-associated proteins are sorted through a series of endosomal compartments for eventual degradation in the lysosome (Lemmon and Traub, 2000). The effects of other "functional knockouts" of the Tsg101/Vps pathway on HIV-1 budding were analyzed. Although Tsg101 is predominantly cytosolic at

steady state (Xie et al., 1998; Zhong et al., 1998), the protein appears to cycle dynamically through the endosomal network. Another protein in the Vps pathway, Vps4, functions in Tsg101 cycling and endosomal trafficking (Babst et al., 1998; Bishop and Woodman, 2000; Bishop and Woodman, 2001). Vps4 is an ATPase of the AAA protein family that appears to utilize the energy derived from ATP hydrolysis to disassemble endosome-associated Vps complexes and thereby allow multiple rounds of sorting. Overexpressing ATPase-defective GFP-Vps4 fusion proteins induces formation of enlarged endosomes (called "class E" compartments in yeast) that are defective in the sorting and recycling of endocytosed substrates (Bishop and Woodman, 2000). The Vps4 mutants also prevent normal Tsg101 trafficking because the protein is trapped on the surface of these aberrant endosomes (Bishop and Woodman, 2001).

384. Vectors expressing ATPase-defective GFP-Vps4 proteins were co-transfected together with the proviral HIV-1 R9 expression vector. Two different dominant negative mutations were used: K173Q, which blocks ATP binding ($Vps4_{K173Q}$), and E228Q, which blocks ATP hydrolysis ($Vps4_{E228Q}$). Protein levels and viral infectivity were again analyzed by Western blotting and in MAGIC assays. As shown in Figs. 4A and 4B, co-expression of the mutant GFP-Vps4 proteins inhibited particle production in a dose-dependent fashion, with release of infectious particles reduced more than 1000-fold at the highest levels of $Vps4_{E228Q}$ tested. Co-transfection with the wild type GFP-Vps4 construct reduced particle production and infectivity only slightly (~2-3-fold). This reduction was reproducible, however, and may reflect stoichiometric imbalances in the Vps pathway arising from Vps4 overexpression. Expression levels for all of the GFP-Vps4 proteins paralleled the levels of transfected DNA (as detected by anti-GFP Western, Fig. 4A, bottom), and none of the Vps4 proteins altered Gag protein synthesis, stability, or cell viability significantly (Fig. 4A, middle). Other controls for this experiment included: 1) mock transfections (lane 1), 2) co-transfection of R9 with a vector expressing GFP only (no effect on budding; lane 2), and 3) transfection of an R9 vector encoding a PTAP to LIRL mutant HIV-1 provirus (arrested budding; lane 3).

385. As the Vps4 mutants interfered with endosomal trafficking, the possibility that particle release might have been inhibited by altered trafficking, activation, or assembly of viral proteins other than Gag (e.g., protease (PR), envelope (Env), or viral protein U) was considered. This was not the case, however, because the mutant GFP-Vps4 proteins also blocked release of viral constructs missing PR or Env and blocked release of Gag-GFP alone (Fig. 4C, lanes 5, 6). As expected, Gag-GFP was released efficiently in the absence of the dominant negative Vps constructs, but was blocked by mutation of the Gag p6 PTAP sequence (lane 3). These experiments demonstrate that the Vps4-induced block to HIV-1 release does not require any other viral proteins, consistent with the idea that the p6 domain of the assembling Gag particle recruits components of the Tsg101/Vps pathway to assist in budding.

(5) A Functional Vps Pathway Is Required for Release of Infectious HIV-1 from Human T Cell Lines

386. It was also tested whether an intact PTAP motif and a functional Vps pathway were required for release of infectious HIV-1 particles from T cells, which are HIV-1 hosts *in vivo*. Although

the transfection efficiencies of T cell lines are low, viral titers produced from MT4 cells 72 h post-transfection were sufficiently high to quantitate using the MAGIC assay. Mutation of HIV-1 p6 PTAP reduced viral titers >300-fold (Fig. 4D) and blocked spreading infections in MT4 cultures (growth curves not shown). Thus, HIV-1 also requires p6 PTAP to replicate in this non-adherent T cell line. Inhibition of the Vps pathway by overexpression of the dominant negative $Vps4_{E228Q}$ mutant also severely reduced infectious HIV-1 titers (>200-fold), whereas overexpression of the wt Vps4 protein had only modest effects (~3-fold reduction). Similar results were obtained in the CEMss T cell line (data not shown). Thus, we conclude that HIV-1 requires both a PTAP motif and a functional Vps pathway to replicate in multiple cell types, including T cells.

(6) A Functional Vps Pathway Is Required for Release of Infectious MLV from 293T Cells

387. The effects of mutant GFP-Vps4 protein overexpression on MLV particle release and infectivity in 293T cells were also tested (Fig. 5). In this case, virion-associated CA release was *blocked*, with a concomitant increase in cytoplasmic levels of the unprocessed Gag protein (lanes 5, 6). Similarly, infectivity was dramatically reduced (>800-fold for $Vps4_{E228Q}$). Thus, the block to viral release imposed by the Vps4 mutants is more general than that imposed by depleting Tsg101, which appears to be HIV-specific. These observations indicate that late endosomal trafficking, which is inhibited by the dominant negative Vps4 proteins, is required for virus release through *both* the PTAP and PPPY motifs.

(7) Tsg101 Depletion and Vps4 Mutations Inhibit Late Stages of Virus Budding

388. Thin section electron microscopy was used to characterize the defects in viral release from Tsg101-depleted and Vps4-inhibited cells. In both experiments, HIV-1 budding arrested with "late domain" phenotypes that were very similar to those observed for control $PTAP^-$ mutant viruses (Fig. 6A). In each case, isolated immature particles remained connected to the plasma membrane via membrane stalks, and budding particles also often formed "clusters" of interconnected virions. Although budding HIV-1 particles were occasionally observed in control cells, these "late domain" mutant phenotypes were rarely seen. The dominant negative Vps4 proteins also arrested MLV particle release at a late stage, with a phenotype that was very similar to that observed for control $PPPY^-$ mutant viruses (Fig. 6B). Thus, HIV-1 and MLV Gag proteins localized to the plasma membrane of both Tsg101-depleted and Vps4-inhibited cells, and appeared to initiate assembly normally. However, the viral particles failed to complete membrane fission, implying that Tsg101 (for HIV-1) and other components of the Vps pathway (for both HIV-1 and MLV) participate in this final stage of viral particle release.

389. Disclosed herein Tsg101 UEV is a multifunctional domain that can simultaneously bind both ubiquitin and the p6 PTAP motif, consistent with Tsg101 detecting or participating in the Ub transfer event required for HIV-1 budding. Disclosed herein the three-dimensional structure of Tsg101 UEV exhibits an E2-like fold, but that the final two helices in the canonical E2 fold are replaced by the PTAP binding site.

b) Experimental Procedures

(1) Yeast Two-hybrid Experiments

390. Yeast two-hybrid reagents and techniques have been reviewed (Bartel and Fields, 1997). Briefly, a bait construct expressing HIV-$1_{NYU/BR5}$ p6 (Gag residues 449-500) fused to the C-terminus of Gal4 DNA binding domain (residues 1-147) was transformed into yeast strain PNY200 (*MATα trp1-901 leu2-3,112 ura3-52 his3-200 ade2 gal4Δ gal80*). Prey constructs expressing cDNA from human spleen poly(A)$^+$ RNA (Clontech) fused to the C-terminus of the Gal4 activation domain (residues 768-881) were transformed into the yeast strain BK100 (*MATa trp1-901 leu2-3,112 ura3-52 his3-200 gal4Δ gal80 LYS2::GAL-HIS3 GAL2-ADE2 met2::GAL7-lacZ*), which incorporates the multiple reporter system (James et al., 1996). PNY200 cells (bait) were mated with BK100 cells (preys) and diploid yeast cells were selected in the presence of 3 mM 3-amino-1,2,4-triazole for the ability to synthesize tryptophan (bait), leucine (prey), histidine (bait/prey interaction), and adenine (bait/prey interaction). Approximately 4.2 million bait/prey pairs were tested, and preys encoding Tsg101 residues 7-390 were isolated twice (Tsg101 NCBI accession # NM_006292).

391. Interactions were confirmed by transforming bait and prey constructs into naïve yeast cells and performing liquid culture β-galactosidase assays. Cultures were grown overnight in synthetic media (-Leu, -Trp, + glucose) in 96 well plates, normalized for optical density, and lysed by addition of 6X lysis/substrate solution in 6X Z-buffer (60mM KCl, 6mM $MgSO_4$, 360mM Na_2HPO_4, 240 mM NaH_2PO_4, 6mg/ml CPRG, 0.12U/ml lyticase, 0.075% NP-40). Cultures were incubated for 2 hr at 37°C, clarified by centrifugation, and the optical absorbance of each supernatant was measured (575 nm). Significant responses were observed for p6 binding to full length Tsg101 (>300-fold) and an N-terminal Tsg101 construct (residues 1-207; >30-fold over background), but not to a C-terminal Tsg101 construct (residues 207-390).

(2) Recombinant Proteins

392. DNA encoding Tsg101 residues 1-145 (the UEV domain) was cloned into pET11d (Novagene), expressed in BL21(DE3) *E. coli* cells, and purified by conventional chromatography. Full details of the purification procedure will be published elsewhere (Pornillos et al., in preparation). Expression of mutant GST-$p6_{NL4-3}$ and GST-Ub proteins followed the procedures described for wt GST-p6 (Jenkins et al., 2001).

(3) Surface Plasmon Resonance

393. Surface plasmon resonance measurements were performed at 20°C using a BIACORE 3000 (Biacore AB, Uppsala, Sweden) equipped with a research-grade CM5 sensor chip. ~10 kRU anti-GST Ab was immobilized on all four flow cells using traditional amine-coupling chemistry (Johnsson et al., 1991). Soluble lysates from *E. coli* expressing GST-p6 proteins were diluted 10X in running buffer (20 mM Na_3PO_4, 150 mM NaCl, 0.005% P20, 50 μg/mL BSA, pH 7.2) and captured individually on three of the antibody surfaces at densities of 1-2.5 kRU. ~2 kRU recombinant GST protein was captured on the fourth (reference) antibody surface.

394. Tsg101 UEV in running buffer was injected in triplicate over the four flow cells at concentrations of 0, 0.09, 0.27, 0.82, 2.5, 7.4, 22, 67, and 200 μM (50 μL/min). Data were collected at a rate of 2 Hz during the 30 s association and dissociation phases. All Tsg 101/p6 interactions reached equilibrium rapidly and dissociated completely within seconds. To obtain equilibrium dissociation constants, the responses at t = 25 s were fit to simple 1:1 interaction binding isotherms (Myszka, 1999).

(4) Plasmid Constructs

395. Vector R9 contains a full-length HIV-1_{NL4-3} expression clone (von Schwedler et al., 1998). Kunkel mutagenesis was used to mutate the HIV-1 Gag p6 PTAP sequence to LIRL without disrupting the overlapping pol reading frame (Huang et al., 1995). R9ΔEnv and R9ΔPR were gifts from Chris Aiken, Vanderbilt University. pGag-GFP contains the rev-independent HIV-1_{HXB2} Gag sequence fused to EGFP, and was a gift from Marilyn Resh (Hermida-Matsumoto and Resh, 2000). The PTAP sequence of Gag p6 was also mutated to LIRL in this construct. The packaging vector pCLeco was used for MLV production (Naviaux et al., 1996). The MLV Gag p12 sequence PPPY was mutated to AAAA (Yuan et al., 1999) by PCR megaprimer mutagenesis of pCLeco. DNA encoding full length Tsg101 was fused to the FLAG coding sequence using PCR and cloned into pIRES2-EGFP (Clontech). Seven silent mutations that render Tsg101-FLAG resistant to siRNA (denoted Tsg*-FLAG in Fig. 2) were introduced into the wild type Tsg101 coding region $_{413}$AACCTCCAGTCTTCTCTCGTC$_{433}$ using Kunkel mutagenesis (mutated nucleotides are underlined). The complete 437aa reading frame of Vps4 was repaired and amplified from ATCC #81449, and cloned into pEGFP-C1 (Clontech) to create a GFP-Vps4 fusion protein. ATPase point mutations K173Q and E228Q were introduced by PCR megaprimer mutagenesis. Cloning details are available upon request.

(5) Small Interfering RNA (siRNA)

396. 21-nt RNA duplexes with symmetric 2-nt 3' (2'-deoxy) thymidine overhangs corresponding to Tsg101 coding nt 413-433 were synthesized and HPLC purified. RNA sequences: Sense, 5' CCU CCA GUC UUC UCU CGU CTT; Antisense, 5' GAC GAG AGA AGA CUG GAG GTT; Inverted sense, 5' CUG CUC UCU UCU GAC CUC CTT; Inverted antisense, GGA GGU CAG AAG AGA GCA GTT. Oligonucleotides were annealed as described (Elbashir et al., 2001).

(6) Protein Expression in 293T and T Cells

397. 293T cells were co-transfected with the indicated amounts of R9 and Vps4 expression vectors in 6 well plates using the calcium phosphate method as described (von Schwedler et al., 1998). All other transfections were performed as described in the text using Lipofectamine2000 (Invitrogen) following the manufacturer's instructions. Cells and supernatants were harvested 24-72h post transfection as described in the text.

(7) Western Blots

398. Cytoplasmic and sucrose-pelleted viral lysates were resolved by SDS-PAGE and blotted for ECL as described (von Schwedler et al., 1998). The following primary antibodies were used: rabbit anti-HIV CA antibody from Hans-Georg Krausslich, Heidelberg, Germany (at 1:2000); rabbit anti-HIV MA

from Didier Trono, Geneva, Switzerland (at 1:50,000); murine monoclonal anti-Tsg101-4A10 from GeneTex, Inc. (at 1:1000), murine monoclonal anti-FLAG M2 from Sigma (at 1:3000), rabbit anti-14-3-3β K19 from Santa Cruz Biotechnology (at 1:3000), anti-GFP-HRP from Clontech (at 1:500), and goat anti-MLV p30 from John Elder, Scripps Institute, La Jolla (at 1:1000).

(8) Viral Replication Assays

399. Infectivity of HIV-1 released into the supernatants from cells transfected with R9 constructs was assayed by MAGIC assay as described (von Schwedler et al., 1998) except that infections were performed at 3 different dilutions in triplicate in 48 well plates. Blue cells and syncytia were counted 2d after infection. MLV infectivity was assayed by packaging pCL-lacZ in the M-MLV retroviral vector pCLeco (Naviaux et al., 1996), transducing NIH3T3 cells, and staining cells for β-galactosidase activity 2 days later, as described (von Schwedler et al., 1998).

(9) Electron Microscopy

400. Transfected 293T cells were fixed with 2.5% glutaraldehyde/ 1% paraformaldehyde for 30min and stained as described (von Schwedler et al., 1998), except that samples were dehydrated in ethanol and embedded in Spurr's plastic.

2. Example 2 Vps4b

401. The human Skd1/Vps4b gene was cloned using standard cloning technology based on the homology of the yeast Skd1 gene. Variant Vps4b nucleic acids were produced using standard recombinant biotechnology to produce the orthologous mutations K180Q and E235Q. These were made as DsRed fusion proteins. The K180Q and E235Q mutations were assayed for their involvement in HIV budding (Figure 7). A Rev-independent GagGFP construct (point cyt anti-CA) was used (Dr. Marylin Reshconstruct). 293T cells were co-transfected with the an expressable Vps4b nucleic acid and an expressable GagGFP construct. Upon co-transfection with the DsRed-Skd1 fusion genes gag budding was blocked by the mutants, but promoted by wild type Vps4b. In immunofluorescence, gag-GFP is more than usually stuck at the membrane.

402. Not only can the Vps4b mutants, K180Q and E235Q block GAG budding they can also block viral budding and viral infectivity. The mutant Vps constructs were cotransfected with HIV virus in cells. Western analysis showed that production of mature GAG proteins, CA and MA was blocked by the K180Q and E235Q mutant Vps4b constructs, as well as the E228Q mutant Vps4a construct (Figure 8A). This indicates that a block in the GAG processing and budding was caused by both VPs4a and Vps4b mutants. Consistent with this result, a MAGIC infectivity analysis showed that the Vps4a and Vps4b mutants blocked production of infectious viral particles as determined by the MAGIC assay (Figure 8B).

403. Vps4a and Vps4b both bind Bc2. GST pulldown experiments showed that Bc2 coprecipitated with both Vps4a and Vps4b (Figure 9A). Overexpression of a DsRed-Bc2 fusion protein along with either cotransfected HIV GAG or HIV virus, showed that production of mature GAG particles

(Figure 9B) and budding of viral particles (Figure 9C). was blocked by the overexpression of the Bc2. Overexpression of cytoplasmic GAG and VPS-25, another yeast ortholog protein did not effect GAG maturation or viral budding (Figure 9B and C). Furthermore, overexpression of Bc2 blocked production of infectious viral particles as determined by a MAGIC assay. (Figure 9D). This is consistent with Bc2 causing a stoichiometric shift in Vps4 associations which prevents Vps4 associations needed for GAG maturation and viral budding.

G. References

404. Amerik, A. Y., Nowak, J., Swaminathan, S., and Hochstrasser, M. (2000). The doa4 deubiquitinating enzyme is functionally linked to the vacuolar protein-sorting and endocytic pathways, Mol Biol Cell *11*, 3365-80.

405. Babst, M., Odorizzi, G., Estepa, E. J., and Emr, S. D. (2000). Mammalian Tumor Susceptibility Gene 101 (TSG101) and the Yeast Homologue, Vps23p, Both Function in Late Endosomal Trafficking, Traffic *1*, 248-258.

406. Babst, M., Wendland, B., Estepa, E. J., and Emr, S. D. (1998). The Vps4p AAA ATPase regulates membrane association of a Vps protein complex required for normal endosome function, Embo J *17*, 2982-93.

407. Bartel, P. L., and Fields, S. (1997). The Yeast Two Hybrid Screen (New York, Oxford University Press, Inc).

408. Berthoud, V. M., Tadros, P. N., and Beyer, E. C. (2000). Connexin and gap junction degradation, Methods *20*, 180-7.

409. Bishop, N., and Woodman, P. (2000). ATPase-defective mammalian Vps4 localizes to aberrant endosomes and impairs cholesterol trafficking, Mol Biol Cell *11*, 227-39.

410. Bishop, N., and Woodman, P. (2001). TSG101/mammalian Vps23 and mammalian Vps28 interact directly and are recruited to Vps4-induced endosomes, J Biol Chem *276*, 11735.

411. Chin, L. S., Raynor, M. C., Wei, X., Chen, H. Q., and Li, L. (2001). Hrs interacts with sorting nexin 1 and regulates degradation of epidermal growth factor receptor, J Biol Chem *276*, 7069-78.

412. Craven, R. C., Harty, R. N., Paragas, J., Palese, P., and Wills, J. W. (1999). Late domain function identified in the vesicular stomatitis virus M protein by use of rhabdovirus-retrovirus chimeras, J Virol *73*, 3359-65.

413. Dunn, R., and Hicke, L. (2001). Domains of the rsp5 ubiquitin-protein ligase required for receptor- mediated and fluid-phase endocytosis, Mol Biol Cell *12*, 421-35.

414. Dupre, S., and Haguenauer-Tsapis, R. (2001). Deubiquitination step in the endocytic pathway of yeast plasma membrane proteins: crucial role of doa4p ubiquitin isopeptidase, Mol Cell Biol *21*, 4482-94.

415. Elbashir, S. M., Harborth, J., Lendeckel, W., Yalcin, A., Weber, K., and Tuschl, T. (2001). Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells, Nature *411*, 494-8.

416. Freed, E. O. (1998). HIV-1 gag proteins: diverse functions in the virus life cycle, Virology *251*, 1-15.

417. Galan, J. M., Wiederkehr, A., Seol, J. H., Haguenauer-Tsapis, R., Deshaies, R. J., Riezman, H., and Peter, M. (2001). Skp1p and the F-box protein Rcy1p form a non-SCF complex involved in recycling of the SNARE Snc1p in yeast, Mol Cell Biol *21*, 3105-17.

418. Garnier, L., Wills, J. W., Verderame, M. F., and Sudol, M. (1996). WW domains and retrovirus budding, Nature *381*, 744-5.

419. Gottlinger, H. G., Dorfman, T., Sodroski, J. G., and Haseltine, W. A. (1991). Effect of mutations affecting the p6 gag protein on human immunodeficiency virus particle release, Proc Natl Acad Sci U S A *88*, 3195-9.

420. Harty, R. N., Brown, M. E., Wang, G., Huibregtse, J., and Hayes, F. P. (2000). A PPxY motif within the VP40 protein of ebola virus interacts physically and functionally with a ubiquitin ligase: implications for filovirus budding, Proc Natl Acad Sci U S A *97*, 13871-6.

421. Harty, R. N., Paragas, J., Sudol, M., and Palese, P. (1999). A proline-rich motif within the matrix protein of vesicular stomatitis virus and rabies virus interacts with WW domains of cellular proteins: implications for viral budding, J Virol *73*, 2921-9.

422. Harvey, K. F., and Kumar, S. (1999). Nedd4-like proteins: an emerging family of ubiquitin-protein ligases implicated in diverse cellular functions, Trends Cell Biol *9*, 166-9.

423. Hermida-Matsumoto, L., and Resh, M. D. (2000). Localization of human immunodeficiency virus type 1 Gag and Env at the plasma membrane by confocal imaging, J Virol *74*, 8670-9.

424. Hicke, L. (1999). Gettin' down with ubiquitin: turning off cell-surface receptors, transporters and channels, Trends Cell Biol *9*, 107-12.

425. Huang, M., Orenstein, J. M., Martin, M. A., and Freed, E. O. (1995). p6Gag is required for particle production from full-length human immunodeficiency virus type 1 molecular clones expressing protease, J Virol *69*, 6810-8.

426. James, P., Halladay, J., and Craig, E. A. (1996). Genomic libraries and a host strain designed for highly efficient two- hybrid selection in yeast, Genetics *144*, 1425-36.

427. Jayakar, H. R., Murti, K. G., and Whitt, M. A. (2000). Mutations in the PPPY motif of vesicular stomatitis virus matrix protein reduce virus budding by inhibiting a late step in virion release, J Virol *74*, 9818-27.

428. Jenkins, Y., Pornillos, O., Rich, R., Myszka, D., Sundquist, W., and Malim, M. (2001). Biochemical Analyses of the Interactions between HIV-1 Vpr and p6Gag, Journal of Virology, in press.

429. Johnsson, B., Lofas, S., and Lindquist, G. (1991). Immobilization of proteins to a carboxymethyldextran-modified gold surface for biospecific interaction analysis in surface plasmon resonance sensors, Anal Biochem *198*, 268-77.

430. Katzmann, D., Babst, M., and Emr, S. D. (2001). Uniquitin-Dependent Sorting into the Multivesicular Body Pathway Requires the Function of a Consrved Endosomal Protein Complex, ESCRT-1, Cell *106*, 145-155.

431. Komada, M., and Kitamura, N. (2001). Hrs and hbp: possible regulators of endocytosis and exocytosis, Biochem Biophys Res Commun *281*, 1065-9.

432. Koonin, E. V., and Abagyan, R. A. (1997). TSG101 may be the prototype of a class of dominant negative ubiquitin regulators, Nat Genet *16*, 330-1.

433. Lemmon, S. K., and Traub, L. M. (2000). Sorting in the endosomal system in yeast and animal cells, Curr Opin Cell Biol *12*, 457-66.

434. Li, L., and Cohen, S. N. (1996). Tsg101: a novel tumor susceptibility gene isolated by controlled homozygous functional knockout of allelic loci in mammalian cells, Cell *85*, 319-29.

435. Li, L., Liao, J., Ruland, J., Mak, T. W., and Cohen, S. N. (2001). A TSG101/MDM2 regulatory loop modulates MDM2 degradation and MDM2/p53 feedback control, Proc Natl Acad Sci U S A *98*, 1619-1624.

436. Li, Y., Kane, T., Tipper, C., Spatrick, P., and Jenness, D. D. (1999). Yeast mutants affecting possible quality control of plasma membrane proteins, Mol Cell Biol *19*, 3588-99.

437. Losko, S., Kopp, F., Kranz, A., and Kolling, R. (2001). Uptake of the ATP-binding cassette (ABC) transporter Ste6 into the yeast vacuole is blocked in the doa4 Mutant, Mol Biol Cell *12*, 1047-59.

438. Myszka, D. G. (1999). Improving biosensor analysis, J Mol Recognit *12*, 279-84.

439. Naviaux, R. K., Costanzi, E., Haas, M., and Verma, I. M. (1996). The pCL vector system: rapid production of helper-free, high-titer, recombinant retroviruses, J Virol *70*, 5701-5.

440. Ott, D. E., Coren, L. V., Chertova, E. N., Gagliardi, T. D., and Schubert, U. (2000). Ubiquitination of HIV-1 and MuLV Gag, Virology *278*, 111-21.

441. Owen, D. J., and Luzio, J. P. (2000). Structural insights into clathrin-mediated endocytosis, Curr Opin Cell Biol *12*, 467-74.

442. Parent, L. J., Bennett, R. P., Craven, R. C., Nelle, T. D., Krishna, N. K., Bowzard, J. B., Wilson, C. B., Puffer, B. A., Montelaro, R. C., and Wills, J. W. (1995). Positionally independent and exchangeable late budding functions of the Rous sarcoma virus and human immunodeficiency virus Gag proteins, J Virol *69*, 5455-60.

443. Patnaik, A., Chau, V., and Wills, J. W. (2000). Ubiquitin is part of the retrovirus budding machinery, Proc Natl Acad Sci U S A *97*, 13069-74.

444. Ponting, C. P., Cai, Y. D., and Bork, P. (1997). The breast cancer gene product TSG101: a regulator of ubiquitination?, J Mol Med *75*, 467-9.

445. Puffer, B. A., Parent, L. J., Wills, J. W., and Montelaro, R. C. (1997). Equine infectious anemia virus utilizes a YXXL motif within the late assembly domain of the Gag p9 protein, J Virol *71*, 6541-6.

446. Puffer, B. A., Watkins, S. C., and Montelaro, R. C. (1998). Equine infectious anemia virus Gag polyprotein late domain specifically recruits cellular AP-2 adapter protein complexes during virion assembly, J Virol *72*, 10218-21.

447. Rotin, D., Staub, O., and Haguenauer-Tsapis, R. (2000). Ubiquitination and endocytosis of plasma membrane proteins: role of Nedd4/Rsp5p family of ubiquitin-protein ligases, J Membr Biol *176*, 1-17.

448. Ruland, J., Sirard, C., Elia, A., MacPherson, D., Wakeham, A., Li, L., Luis De La Pompa, J., Cohen, S. N., and Mak, T. W. (2001). p53 Accumulation, defective cell proliferation, and early embryonic lethality in mice lacking tsg101, Proc Natl Acad Sci U S A *98*, 1859-1864.

449. Schubert, U., Anton, L. C., Gibbs, J., Norbury, C. C., Yewdell, J. W., and Bennink, J. R. (2000a). Rapid degradation of a large fraction of newly synthesized proteins by proteasomes, Nature *404*, 770-4.

450. Schubert, U., Ott, D. E., Chertova, E. N., Welker, R., Tessmer, U., Princiotta, M. F., Bennink, J. R., Krausslich, H. G., and Yewdell, J. W. (2000b). Proteasome inhibition interferes with gag polyprotein processing, release, and maturation of HIV-1 and HIV-2, Proc Natl Acad Sci U S A *97*, 13057-62.

451. Shih, S. C., Sloper-Mould, K. E., and Hicke, L. (2000). Monoubiquitin carries a novel internalization signal that is appended to activated receptors, Embo J *19*, 187-98.

452. Strack, B., Calistri, A., Accola, M. A., Palu, G., and Gottlinger, H. G. (2000). A role for ubiquitin ligase recruitment in retrovirus release, Proc Natl Acad Sci U S A *97*, 13063-8.

453. VanDemark, A. P., Hofmann, R. M., Tsui, C., Pickart, C. M., and Wolberger, C. (2001). Molecular Insights into Polyubiquitin Chain Assembly. Crystal Structure of the Mms2/Ubc13 Heterodimer, Cell *105*, 711-20.

454. VerPlank, L., Bouamr, F., LaGrassa, T. J., Agresta, B., Kikonyogo, A., Leis, J., and Carter, C. A. (2001). Tsg101, a homologue of ubiquitin-conjugating (E2) enzymes, binds the L domain in HIV type 1 Pr55Gag, Proc Natl Acad Sci U S A *98*, 7724-9.

455. Vogt, V. M. (2000). Ubiquitin in retrovirus assembly: actor or bystander?, Proc Natl Acad Sci U S A *97*, 12945-7.

456. von Schwedler, U. K., Stemmler, T. L., Klishko, V. Y., Li, S., Albertine, K. H., Davis, D. R., and Sundquist, W. I. (1998). Proteolytic refolding of the HIV-1 capsid protein amino-terminus facilitates viral core assembly, Embo J *17*, 1555-68.

457. Wang, G., McCaffery, J. M., Wendland, B., Dupre, S., Haguenauer-Tsapis, R., and Huibregtse, J. M. (2001). Localization of the Rsp5p Ubiquitin-Protein Ligase at Multiple Sites within the Endocytic Pathway, Mol Cell Biol *21*, 3564-75.

458. Wills, J. W., and Craven, R. C. (1991). Form, function, and use of retroviral gag proteins, Aids *5*, 639-54.

459. Xie, W., Li, L., and Cohen, S. N. (1998). Cell cycle-dependent subcellular localization of the TSG101 protein and mitotic and nuclear abnormalities associated with TSG101 deficiency, Proc Natl Acad Sci U S A *95*, 1595-600.

460. Yuan, B., Campbell, S., Bacharach, E., Rein, A., and Goff, S. P. (2000). Infectivity of moloney murine leukemia virus defective in late assembly events is restored by late assembly domains of other retroviruses, J Virol *74*, 7250-60.

461. Yuan, B., Li, X., and Goff, S. P. (1999). Mutations altering the moloney murine leukemia virus p12 Gag protein affect virion production and early events of the virus life cycle, Embo J *18*, 4700-10.

462. Zhong, Q., Chen, Y., Jones, D., and Lee, W. H. (1998). Perturbation of TSG101 protein affects cell cycle progression, Cancer Res *58*, 2699-702.

EN.REFLIST

H. Sequences

1. **SEQ ID NO:1 TSG101 protein sequence one example Genbank accession number Q99816.**
2. **SEQ ID NO:2 DNA for TSG101 NM_006292. human TSG101 gene**
3. **SEQ ID NO:3 TSG101 protein encoded by SEQ ID NO:2**
4. **SEQ ID NO:4. Vps4A protein Genbank accession number AAG01470.**
5. **SEQ ID NO:5: Vps4A Protein Genbank accession number NP_037377**
6. **SEQ ID NO:6 Vps4B Protein Genbank Acession Number AAG01471.**
7. **SEQ ID NO:7 Vps4 Mouse Genbank Accession No. NP_ 569053**
8. **SEQ ID NO:8 Vps4B gene Genbank Accession Number AF282904**
9. **SEQ ID NO:9 Vps4B protein encoded by SEQ ID NO:8 AF282904**
10. **SEQ ID NO:10 Vps4A human gene Genbank accession number AF282903**
11. **SEQ ID NO:11 Vps4A Protein encoded by SEQ ID NO:10 AF282903**

VI. CLAIMS

What is claimed is:

1. A method for reducing interactions between TSG101 and Vps4A, Vps4B, Ub, or Bc2 , comprising incubating an inhibitor of the interaction between TSG101 and Vps4A, Vps4B, Ub, or Bc2.

2. A method for inhibiting HIV budding comprising administering an inhibitor of the interaction between TSG101 and Vps4A, Vps4B, Ub, or Bc2 .

3. A method of treating a subject comprising administering to the subject an inhibitor of HIV budding, wherein the inhibitor reduces the interaction between TSG101 and Vps4A, Vps4B, Ub, or Bc2 , and wherein the subject is in need of such treatment.

4. A method of identifying an inhibitor of an interaction between TSG101 and Vps4A, Vps4B, Ub, or Bc2 comprising incubating a library of molecules with TSG101 forming a mixture, and identifying the molecules that disrupt the interaction between TSG101 and Vps4A, Vps4B, Ub, or Bc2, wherein the interaction disrupted comprises an interaction between the Vps4A, Vps4B, Ub, or Bc2 and an amino acid of TSG101.

5. The method of claim 5, wherein the step of isolating comprises incubating the mixture with molecule comprising a Vps4A, Vps4B, Ub, or Bc2 variant or fragment thereof.

6. A method of identifying an inhibitor of an interaction between TSG101 and Vps4A, Vps4B, Ub, or Bc2 comprising incubating a library of molecules with Vps4A, Vps4B, Ub, or Bc2 forming a mixture, and identifying the molecules that disrupt the interaction between Vps4A, Vps4B, Ub, or Bc2 and TSG101, wherein the interaction disrupted comprises an interaction between the Vps4A, Vps4B, Ub, or Bc2 and an amino acid of TSG101.

7. The method of claim 6, wherein the step of isolating comprises incubating the mixture with molecule comprising a TSG 101 UEV domain.

8. A composition identified by the process of claims 4 or 6.

9. A composition capable of being identified by the process of claim 4 or 6.

10. The method of claims 4 or 6, further comprising incubating the mixture with at least one of Vps4A, Vps4B, Ub, or Bc2.

11. A method of manufacturing a composition for inhibiting the interaction between TSG101 and Vps4A, Vps4B, Ub, or Bc2 comprising synthesizing the inhibitor of claims 4 or 6.

12. The method of claim 11 further comprising mixing a pharmaceutical carrier with the inhibitor.

13. A method of manufacturing a composition for inhibiting the interaction between TSG101 and GAG comprising admixing the inhibitor with a pharmaceutical carrier.

14. The method of claims 1-3 wherein the inhibitor prevents expression of TSG101 mRNA.

15. The method of claims 1-3, wherein the inhibitor interacts with the mRNA of TSG101.

16. The method of claim 1-3, wherein the inhibitor is an interfering RNA.

18. The method of claim 1-3, wherein the interfering RNA has the sequence set forth in SEQ ID NO:3.

19. A method for inhibiting viral budding in a virus comprising proteins that interact with TSG101, comprising administering an inhibitor of the Vps pathway in a cell.

20. The method of claim 19, wherein the virus is HIV or Ebola.

21. The method of claim 19, wherein the inhibitor prevents normal trafficking of TSG101.

22. The method of claim 19, wherein the inhibitor reduces the release of TSG101 from endosomes.

23. The method of claim 19, wherein the inhibitor prevents normal Vps4 function.

24. The method of claim 19, wherein the inhibitor interacts with Vps4.

25. A method for reducing TSG101 function, comprising incubating an inhibitor of the kinase activity of Vps4A or Vps4B.

26. A method for inhibiting HIV budding comprising administering an inhibitor of the kinase activity of Vps4A or Vps4B.

27. A method of treating a subject comprising administering to the subject an inhibitor of HIV budding, wherein the inhibitor reduces the kinase activity of Vps4AorVps4B, and wherein the subject is in need of such treatment.

28. A method of identifying an inhibitor of Vps4A or 4B ATPase activity comprising, a) administering a composition to a system, wherein the system supports Vps4A or 4B ATPase activity, b) assaying the effect of the composition on the amount of Vps4A or 4B ATPase activity in the system, and c) selecting a composition which causes a decrease in the amount of Vps4A or 4B ATPase activity present in the system relative to the system without the addition of the composition.

29. The method of claim 28, wherein the system comprises retroviral budding activity and the Vps4A or 4B ATPase activity is determined by assaying the retroviral budding activity.

30. A method of identifying an inhibitor of HIV budding comprising, a) administering a composition to a system, wherein the system supports HIV budding via a Vps4A or 4B ATPase activity, b) assaying the effect of the composition on the amount of HIV budding in the system, and c) selecting a composition which causes a decrease in the amount of HIV budding present in the system because of an inhibition of the Vps4A or 4B ATPase activity relative to the system without the addition of the composition.

31. A method of inhibiting HIV budding comprising administering a composition, wherein the composition prevents HIV budding, wherein the composition is defined as a composition capable of being identified by administering the composition to a system, wherein the system supports HIV-budding via a Vps4A or 4B ATPase activity interaction, assaying the effect of the composition on the amount of HIV budding in the system, and selecting a composition which causes a decrease in the amount of HIV budding present in the system because of an inhibition of the Vps4A or 4B ATPase activity relative to the system

without the addition of the composition.

32. A method of inhibiting HIV budding comprising administering a composition that reduces Vps4A or 4B ATPase activity.

33. A method of making a composition capable of inhibiting HIV budding comprising admixing a compound with a pharmaceutically acceptable carrier, wherein the compound is identified by administering the compound to a system, wherein the system supports HIV budding via a Vps4A or 4B ATPase activity, assaying the effect of the compound on the amount of HIV budding in the system, and selecting a compound which causes a decrease in the amount of HIV budding in the system because of an inhibition of the Vps4A or 4B ATPase activity, relative to the system without the addition of the compound.

34. A method of manufacturing an inhibitor to HIV budding comprising, a) administering a composition to a system, wherein the system supports HIV budding via a Vps4A or 4B ATPase activity, b) assaying the effect of the composition on the amount of HIV budding in the system, c) selecting a composition which causes a decrease in the amount of HIV budding present in the system because of an inhibition of the Vps4A or 4B ATPase activity, relative to the system with the addition of the composition, and d) synthesizing the composition.

35. The method of claim 34, further comprising the step of admixing the composition with a pharmaceutical carrier.

36. A method of identifying an inhibitor of Vps4A or 4B ATPase activity comprising a) administering a composition to a system, wherein the system comprises Vps4A or 4B ATPase activity, b) assaying the effect of the composition on a Vps4A or 4B ATPase activity, and c) selecting a composition which inhibits a Vps4A or 4B ATPase activity.

37. A cells comprising, a) a regulatable nucleic acid comprising sequence encoding a Vps4A or 4B gene and b) a nucleic acid comprising sequence encoding a budding retrovirus as well as cells further comprising a regulatable nucleic acid comprising sequence encoding a TSG101 gene.

38. Disclosed are cells that further comprising an inhibitor of a Vps4A or 4B ATPase activity.

39. A method of inhibiting HIV budding comprising, administering an dominant negative mutant of Vps4A, wherein the dominant negative mutant substitutes a Q at position 228 of SEQ ID NO:11.

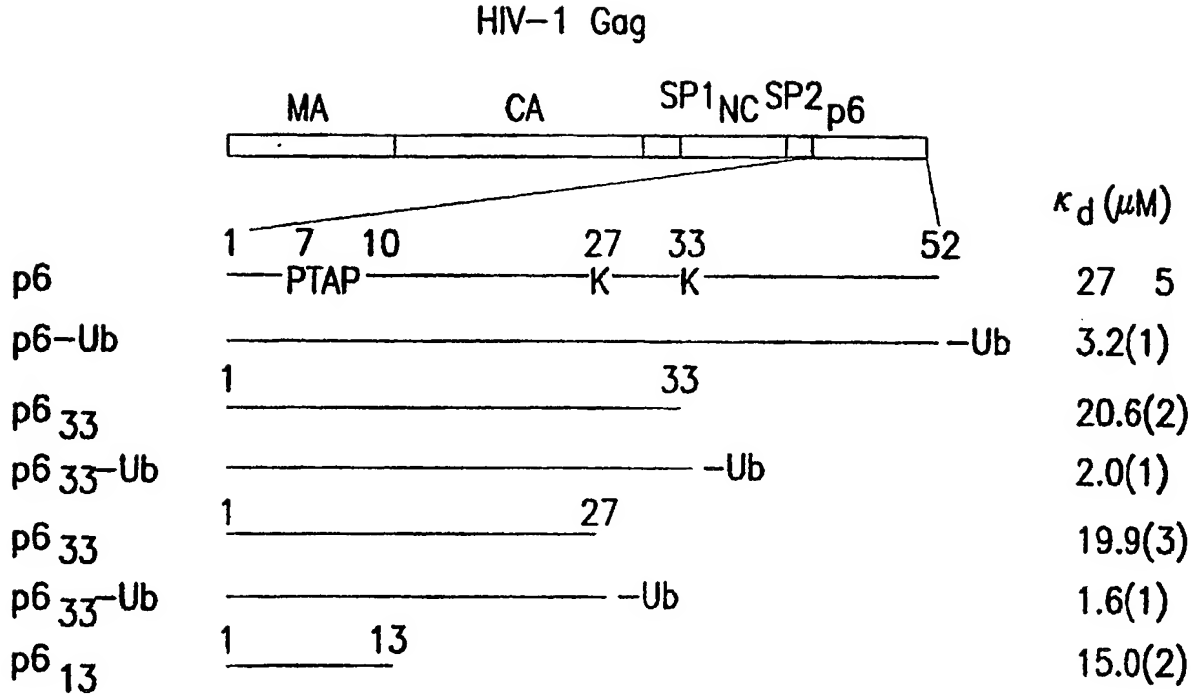

FIG. 1A

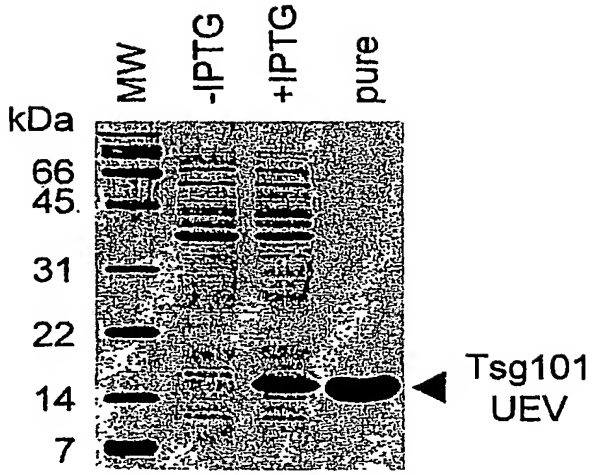

FIG. 1B

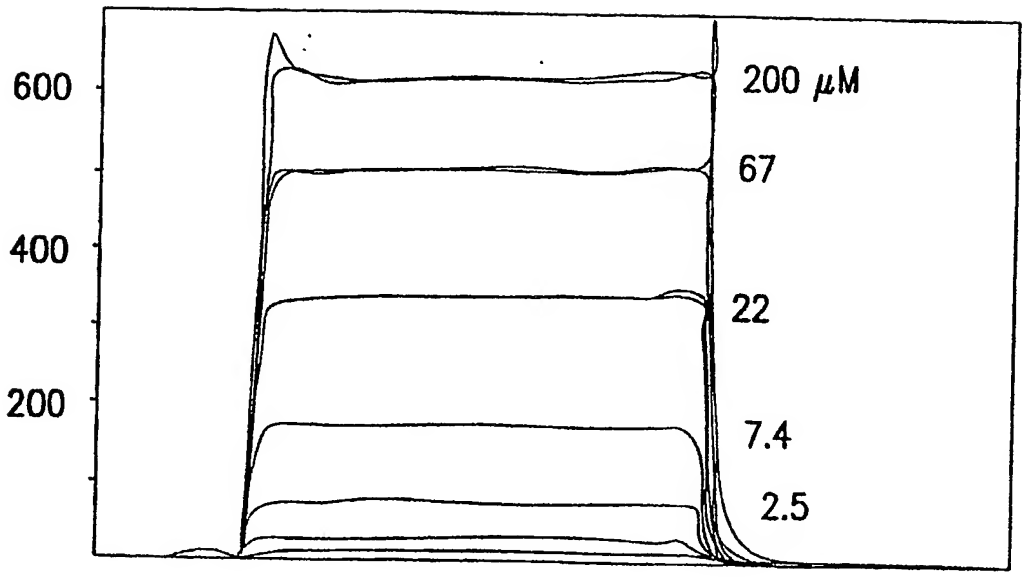

FIG. 1C

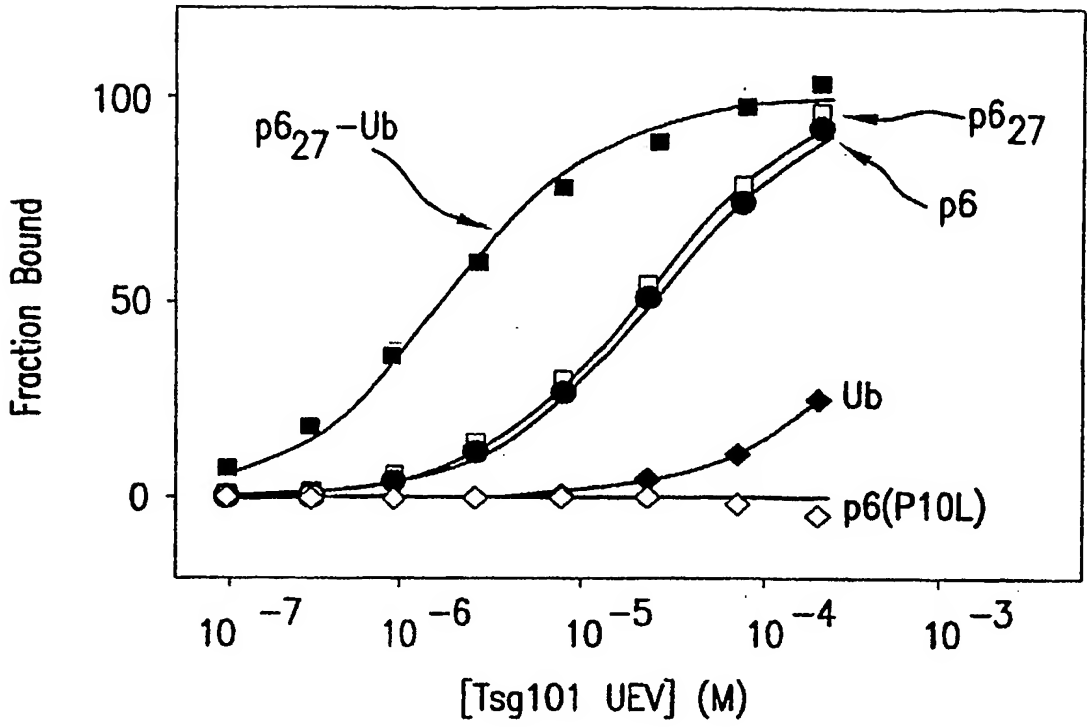

FIG. 1D

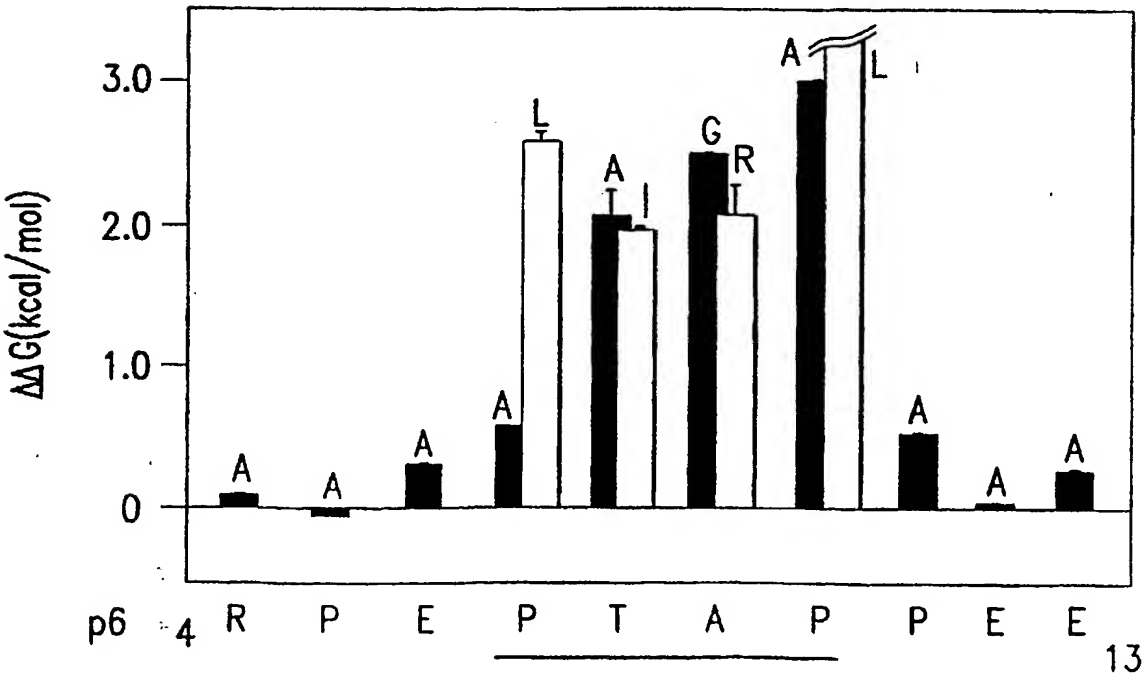

FIG. 1E

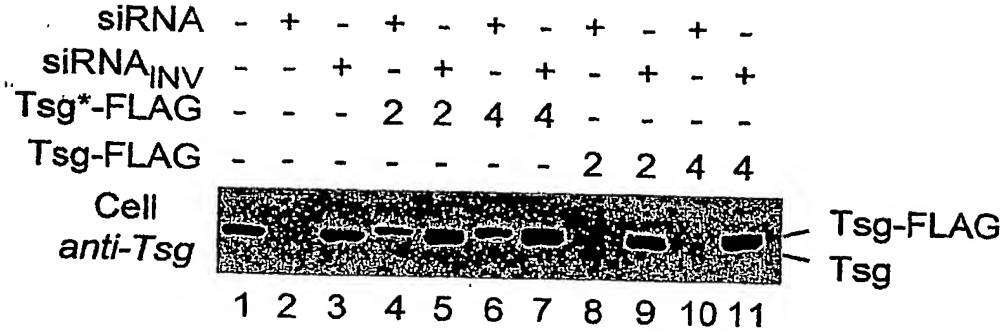
siRNA - + - + - + - + - + -
siRNA$_{INV}$ - - + - + - + - + - +
Tsg*-FLAG - - - 2 2 4 4 - - - -
Tsg-FLAG - - - - - - - 2 2 4 4
Cell
anti-Tsg
Tsg-FLAG
Tsg
1 2 3 4 5 6 7 8 9 10 11

FIG.2A

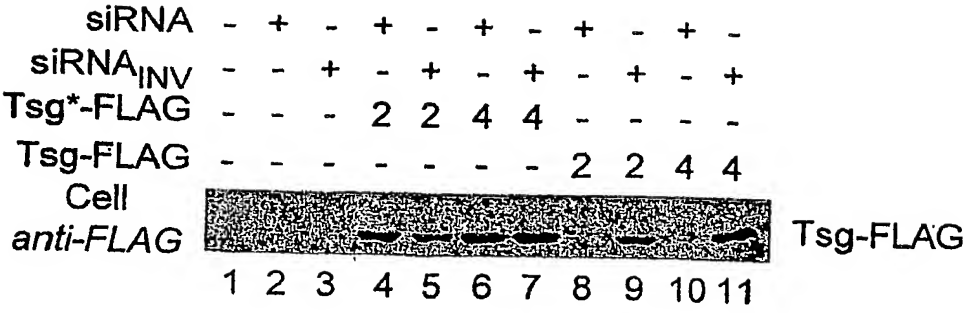
siRNA - + - + - + - + - + -
siRNA$_{INV}$ - - + - + - + - + - +
Tsg*-FLAG - - - 2 2 4 4 - - - -
Tsg-FLAG - - - - - - - 2 2 4 4
Cell
anti-FLAG
Tsg-FLAG
1 2 3 4 5 6 7 8 9 10 11

FIG.2B

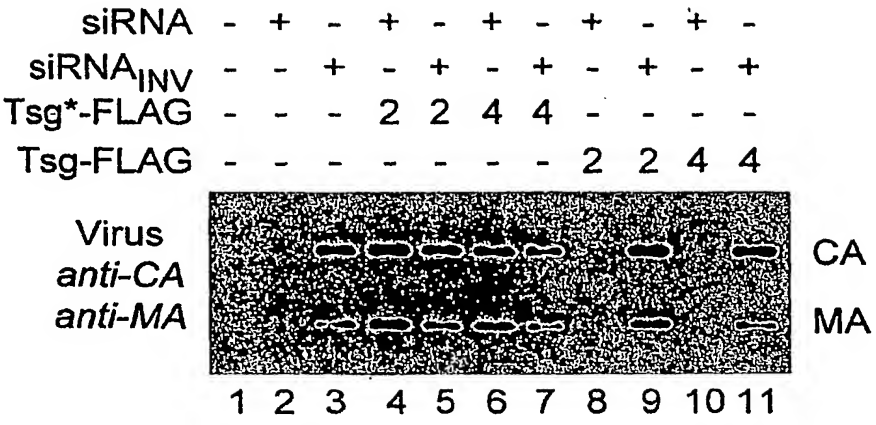

FIG.2C

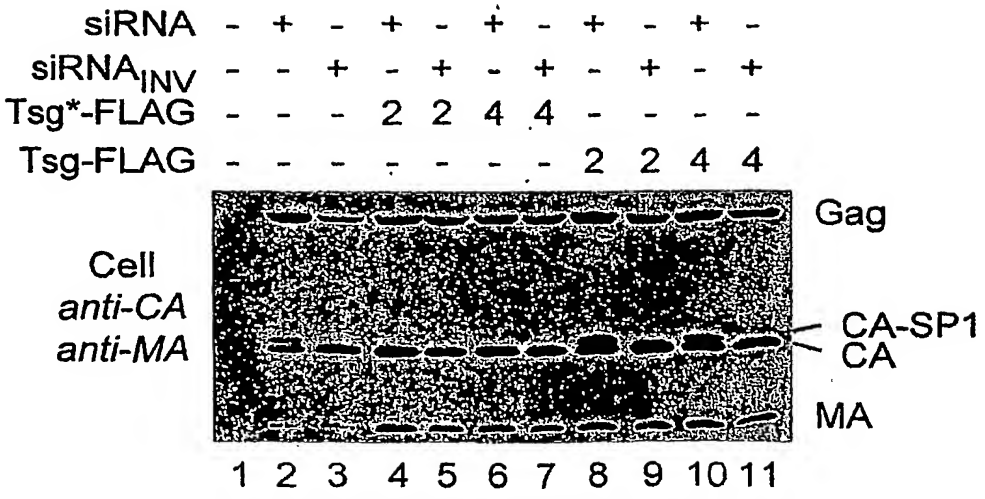

FIG.2D

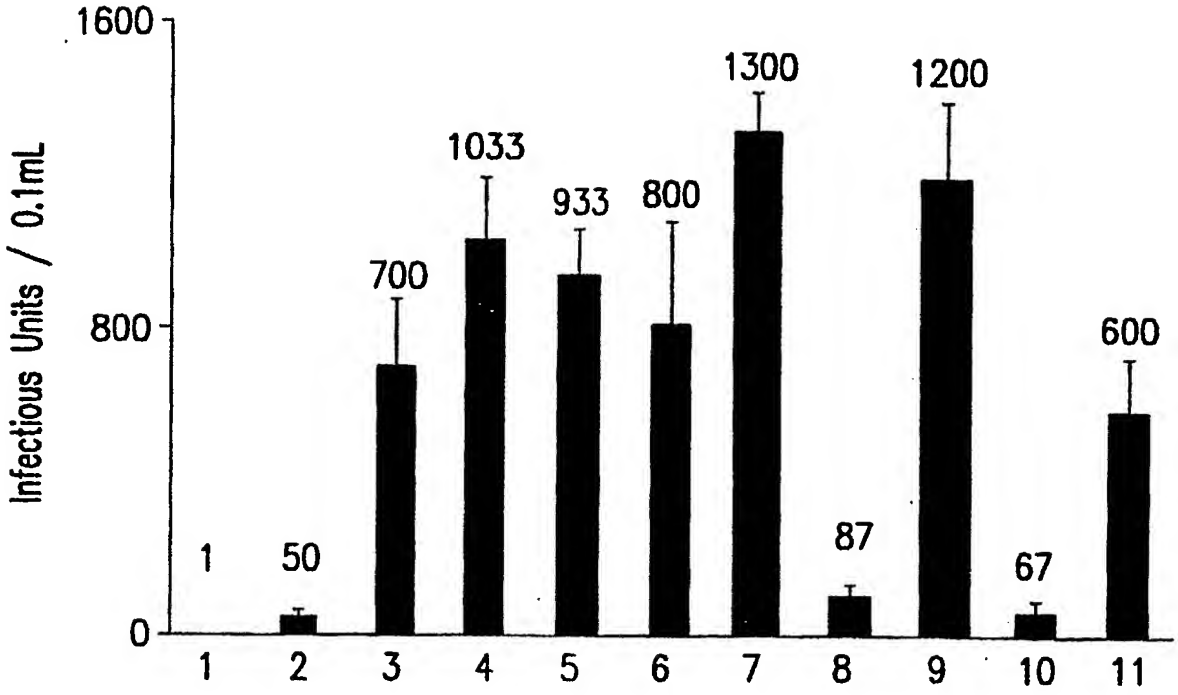

FIG.2E

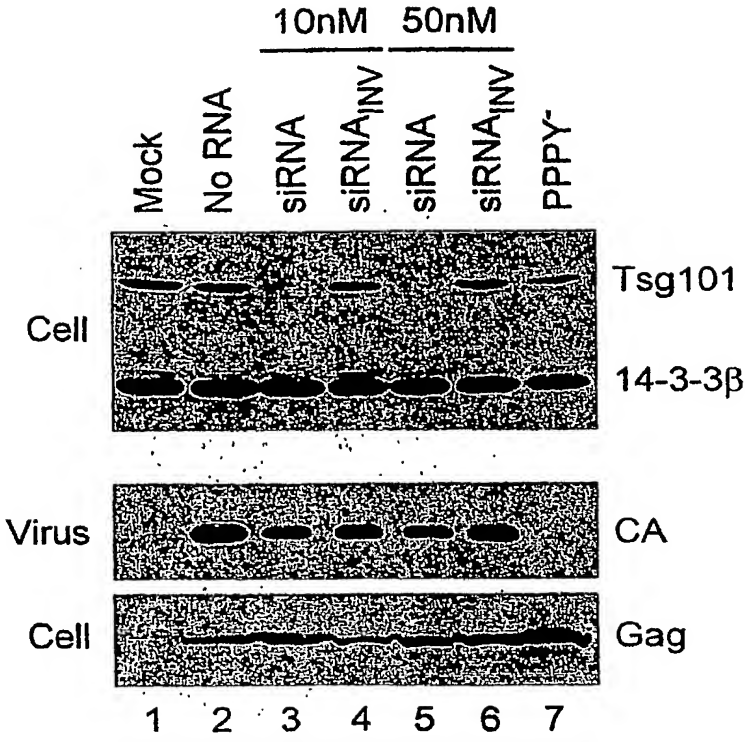

FIG.3A

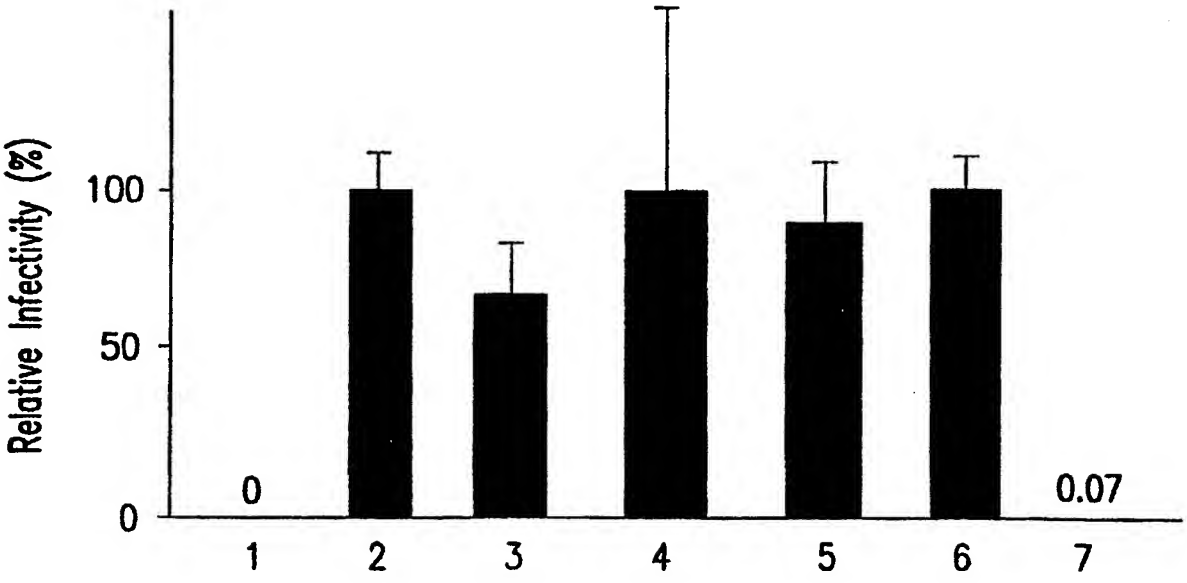

FIG.3B

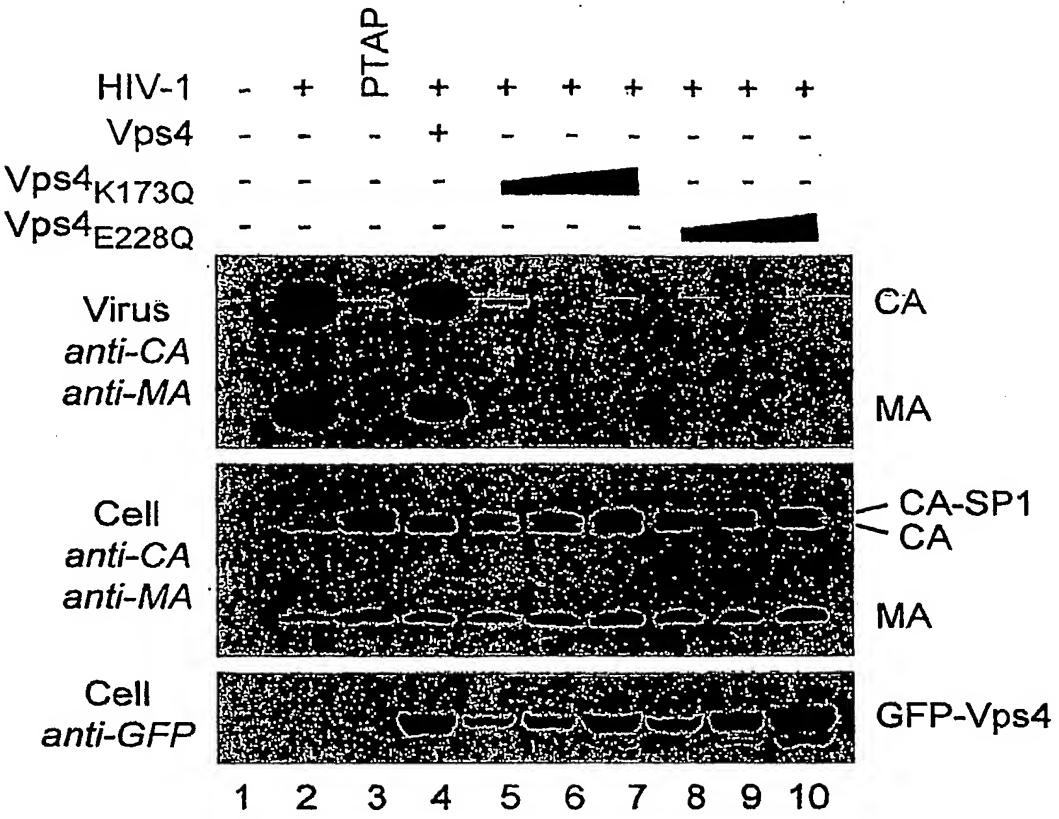

FIG.4A

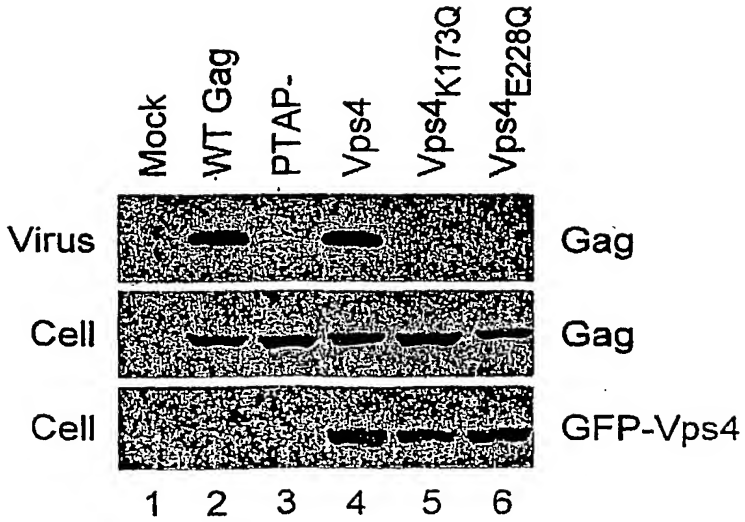

FIG.4C

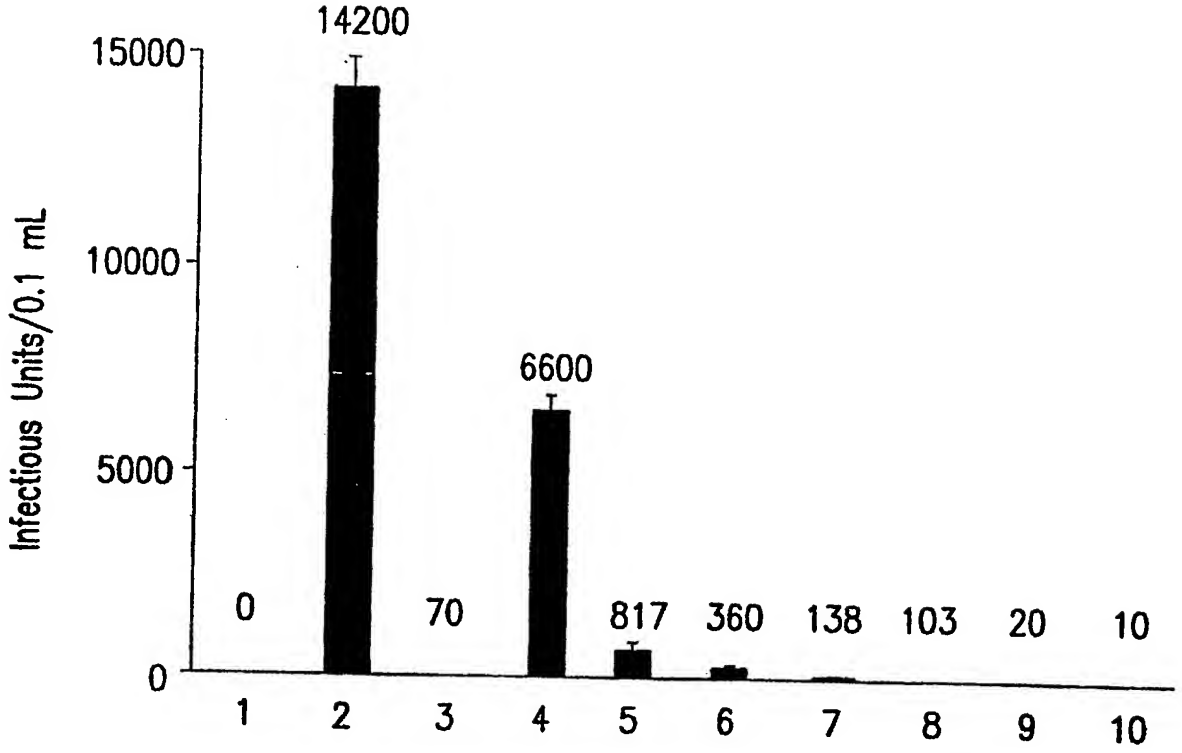
Infectious Units/0.1 mL
15000
10000
5000
0
14200
6600
0
70
817
360
138
103
20
10
1
2
3
4
5
6
7
8
9
10

FIG.4B

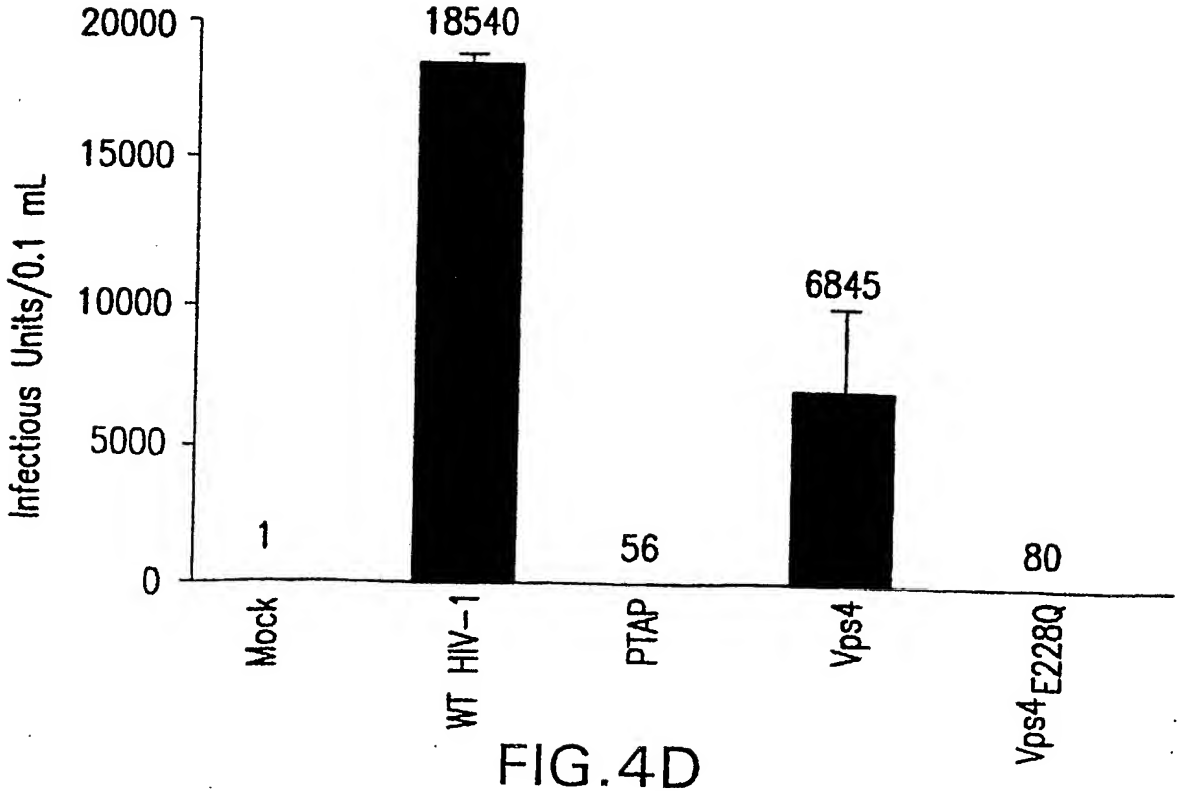
Infectious Units/0.1 mL
20000
15000
10000
5000
0
18540
6845
1
56
80
Mock
WT HIV-1
PTAP
Vps4
Vps4E228Q

FIG.4D

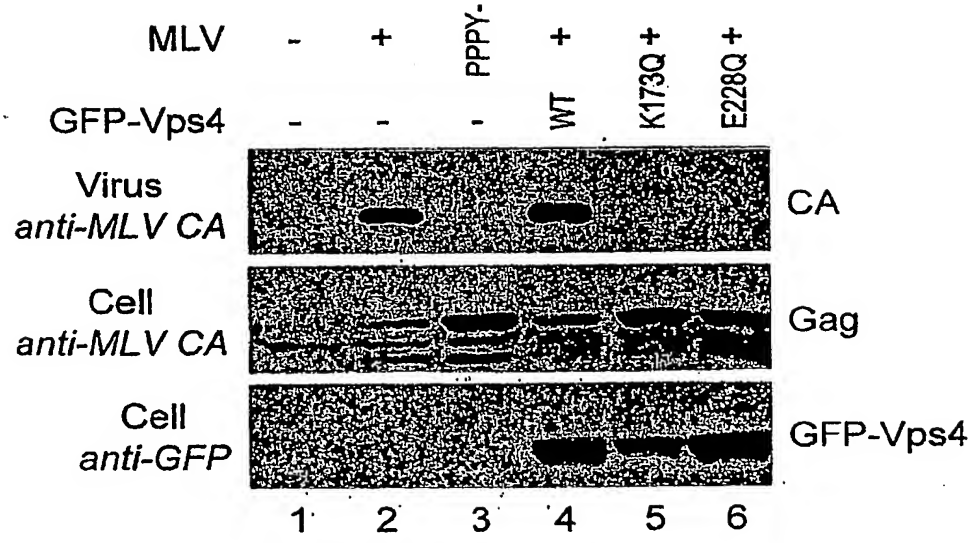

FIG.5A

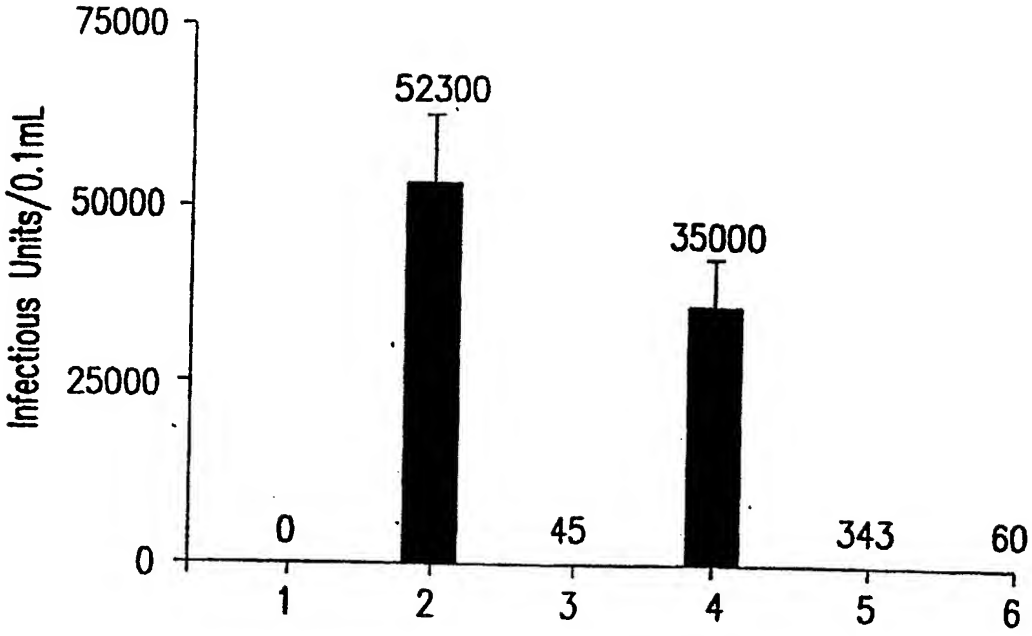

FIG.5B

FIG.6A

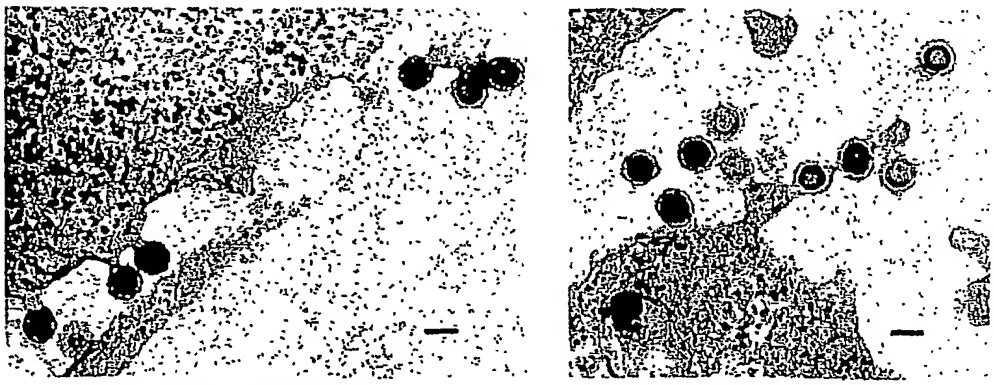

FIG.6B

FIG. 7B

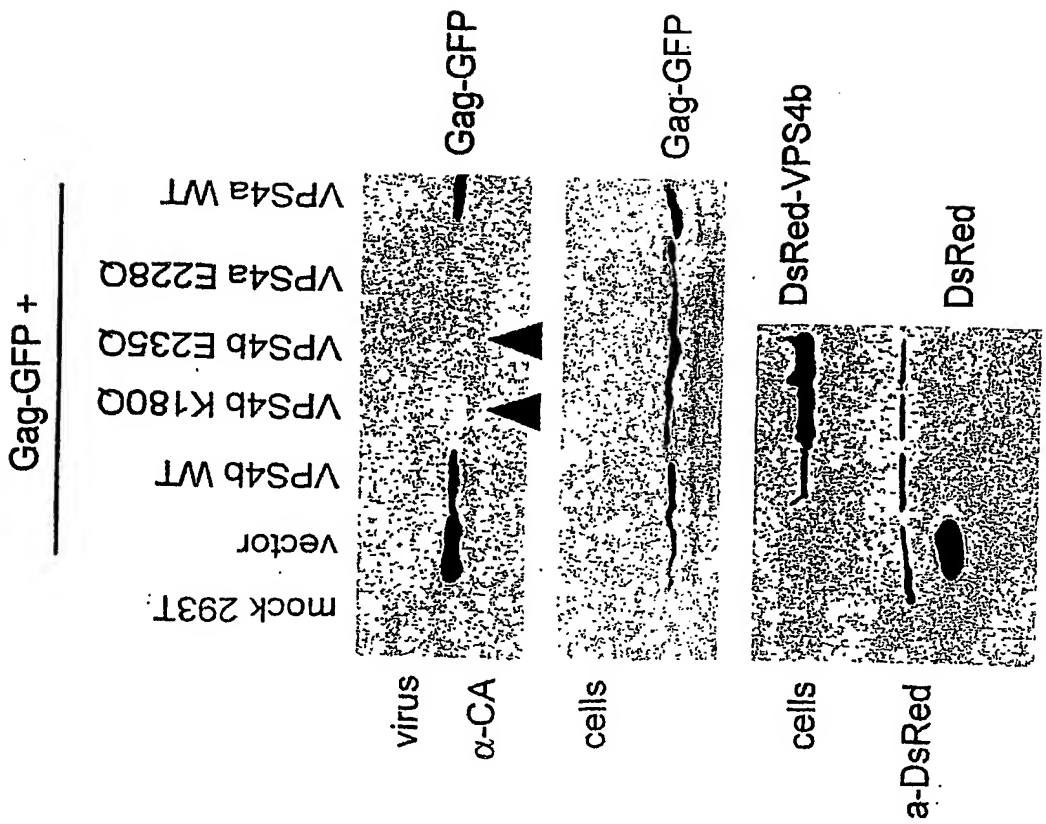

FIG. 7A

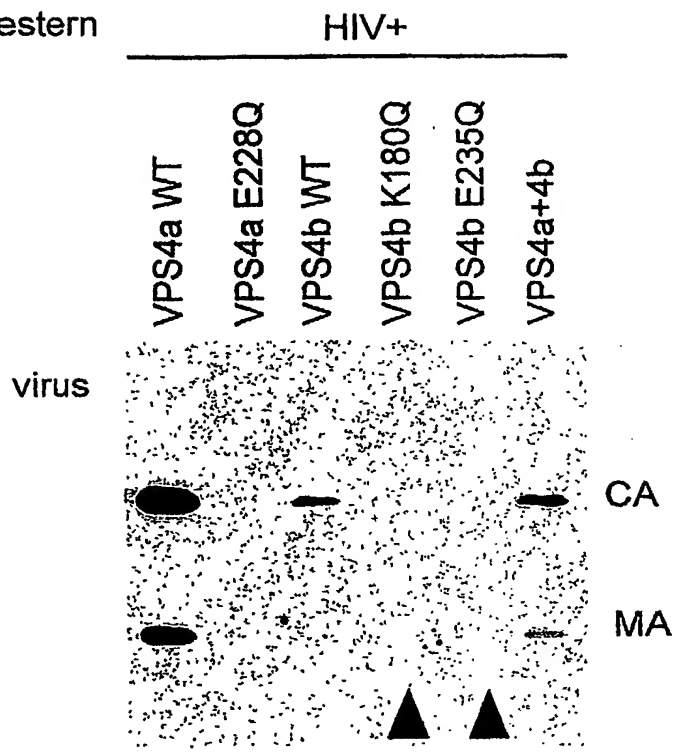

FIG.8A

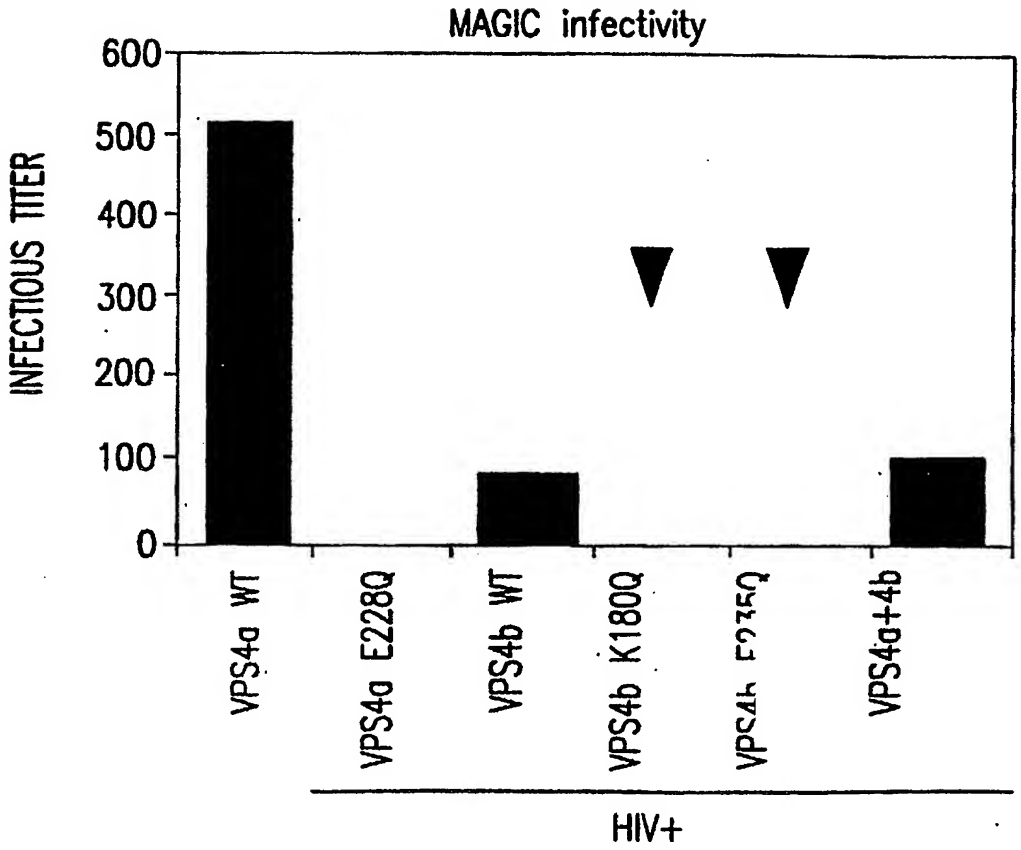

FIG.8B

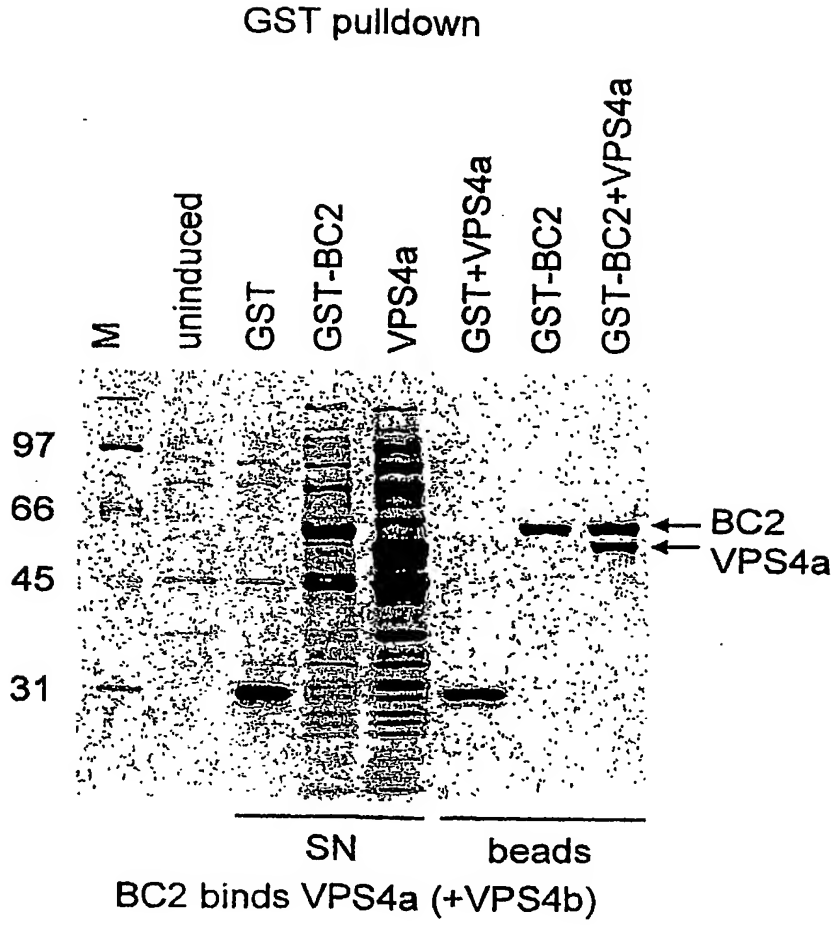

FIG.9A

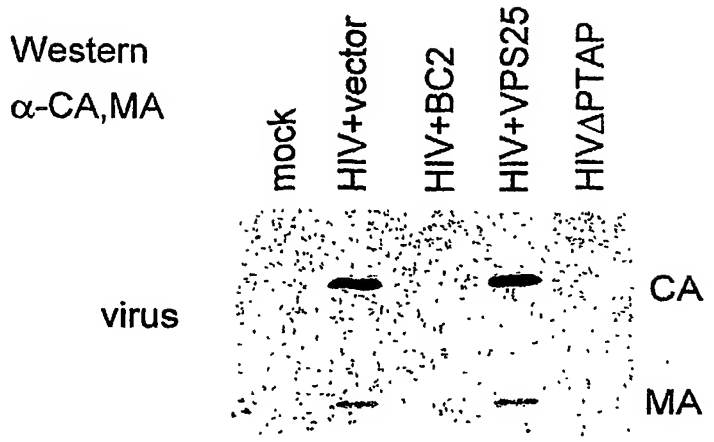

FIG.9B

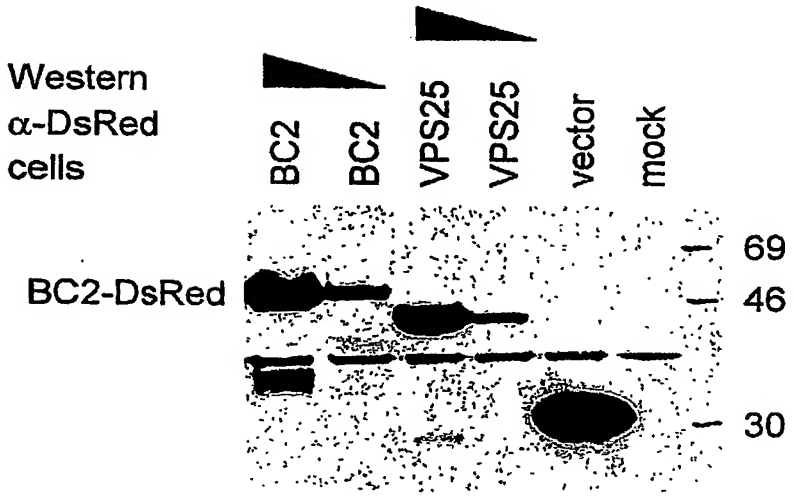

FIG.9C

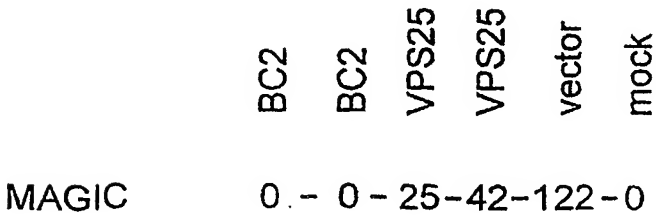

FIG.9D

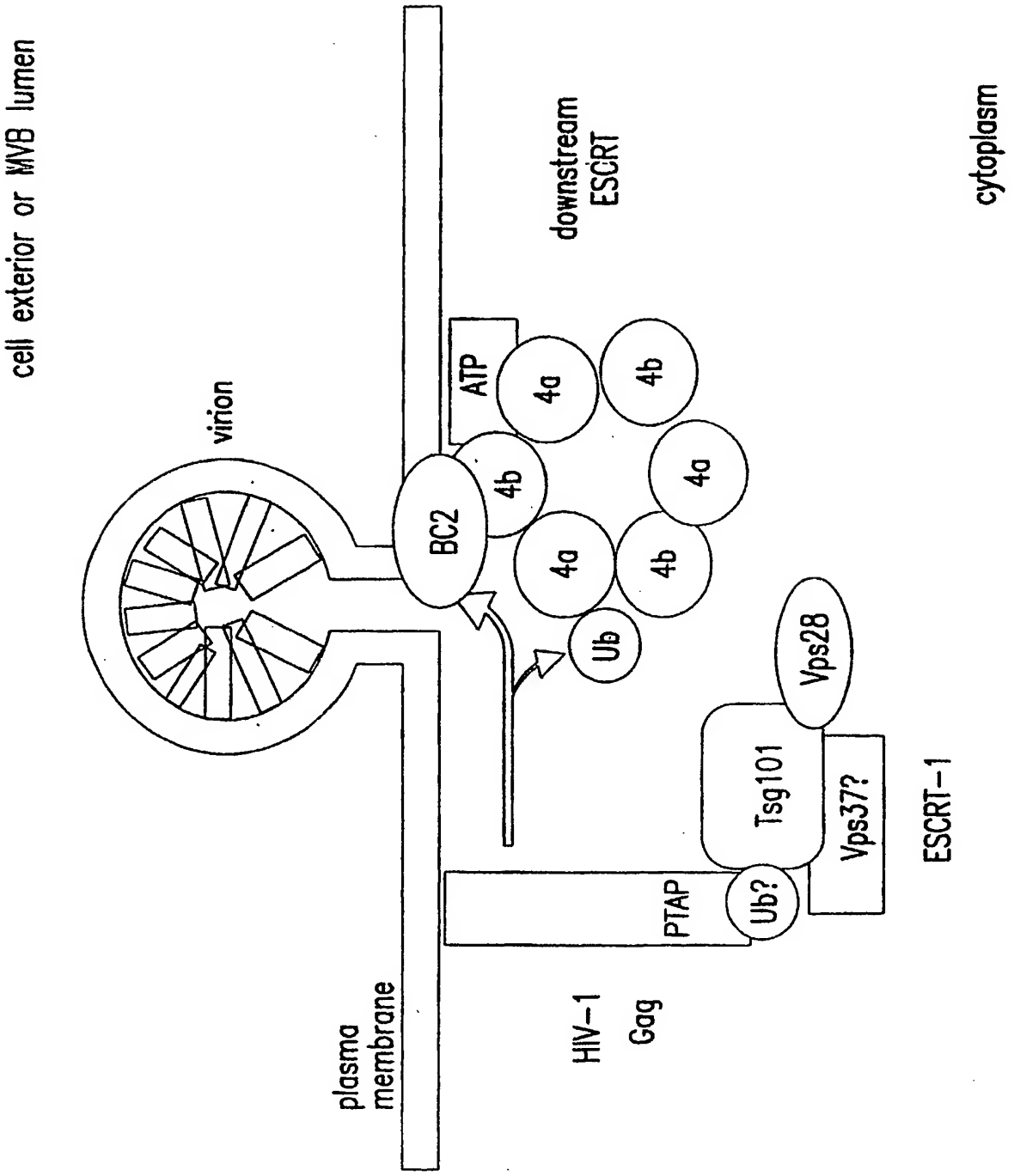

FIG.10

SEQUENCE LISTING

<110> University of Utah Research Foundation
Sundquist, Wesley I.
Garrus, Jennifer E.
von Schwedler, Uta K.

<120> HUMAN VPS4A AND VPS4B FUNCTIONS IN RETROVIRAL BUDDING

<130> 21101.0029P1

<150> 60/382,702
<151> 2003-05-21

<150> 60/381,027
<151> 2002-05-15

<160> 11

<170> FastSEQ for Windows Version 4.0

<210> 1
<211> 390
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:/note = synthetic construct

<400> 1

```
Met Ala Val Ser Glu Ser Gln Leu Lys Lys Met Val Ser Lys Tyr Lys
 1               5                  10                  15
Tyr Arg Asp Leu Thr Val Arg Glu Thr Val Asn Val Ile Thr Leu Tyr
            20                  25                  30
Lys Asp Leu Lys Pro Val Leu Asp Ser Tyr Val Phe Asn Asp Gly Ser
        35                  40                  45
Ser Arg Glu Leu Met Asn Leu Thr Gly Thr Ile Pro Val Pro Tyr Arg
    50                  55                  60
Gly Asn Thr Tyr Asn Ile Pro Ile Cys Leu Trp Leu Leu Asp Thr Tyr
65                  70                  75                  80
Pro Tyr Asn Pro Pro Ile Cys Phe Val Lys Pro Thr Ser Ser Met Thr
                85                  90                  95
Ile Lys Thr Gly Lys His Val Asp Ala Asn Gly Lys Ile Tyr Leu Pro
            100                 105                 110
Tyr Leu His Glu Trp Lys His Pro Gln Ser Asp Leu Leu Gly Leu Ile
        115                 120                 125
Gln Val Met Ile Val Val Phe Gly Asp Glu Pro Pro Val Phe Ser Arg
    130                 135                 140
Pro Ile Ser Ala Ser Tyr Pro Pro Tyr Gln Ala Thr Gly Pro Pro Asn
145                 150                 155                 160
Thr Ser Tyr Met Pro Gly Met Pro Gly Gly Ile Ser Pro Tyr Pro Ser
                165                 170                 175
Gly Tyr Pro Pro Asn Pro Ser Gly Tyr Pro Gly Cys Pro Tyr Pro Pro
            180                 185                 190
Gly Gly Pro Tyr Pro Ala Thr Thr Ser Ser Gln Tyr Pro Ser Gln Pro
        195                 200                 205
```

```
Pro Val Thr Thr Val Gly Pro Ser Arg Asp Gly Thr Ile Ser Glu Asp
    210                 215                 220
Thr Ile Arg Ala Ser Leu Ile Ser Ala Val Ser Asp Lys Leu Arg Trp
225                 230                 235                 240
Arg Met Lys Glu Glu Met Asp Arg Ala Gln Ala Glu Leu Asn Ala Leu
                245                 250                 255
Lys Arg Thr Glu Glu Asp Leu Lys Lys Gly His Gln Lys Leu Glu Glu
            260                 265                 270
Met Val Thr Arg Leu Asp Gln Glu Val Ala Glu Val Asp Lys Asn Ile
        275                 280                 285
Glu Leu Leu Lys Lys Lys Asp Glu Glu Leu Ser Ser Ala Leu Glu Lys
    290                 295                 300
Met Glu Asn Gln Ser Glu Asn Asn Asp Ile Asp Glu Val Ile Ile Pro
305                 310                 315                 320
Thr Ala Pro Leu Tyr Lys Gln Ile Leu Asn Leu Tyr Ala Glu Glu Asn
                325                 330                 335
Ala Ile Glu Asp Thr Ile Phe Tyr Leu Gly Glu Ala Leu Arg Arg Gly
            340                 345                 350
Val Ile Asp Leu Asp Val Phe Leu Lys His Val Arg Leu Leu Ser Arg
        355                 360                 365
Lys Gln Phe Gln Leu Arg Ala Leu Met Gln Lys Ala Arg Lys Thr Ala
    370                 375                 380
Gly Leu Ser Asp Leu Tyr
385                 390
```

<210> 2
<211> 1550
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:/note =
synthetic construct

<400> 2

```
gtagtggtgc cgacttcctg ttgtttgagg ccgggttggg ggtgtgcgat tgtgtgggac      60
ggtctggggc agcccagcag cggctgaccc tctgcctgcg gggaagggag tcgccaggcg     120
gccgtcatgg cggtgtcgga gagccagctc aagaaaatgg tgtccaagta caaatacaga     180
gacctaactg tacgtgaaac tgtcaatgtt attactctat acaaagatct caaacctgtt     240
ttggattcat atgtttttaa cgatggcagt tccagggaac taatgaacct cactggaaca     300
atccctgtgc cttatagagg taatacatac aatattccaa tatgcctatg gctactggac     360
acatacccat ataatccccc tatctgtttt gttaagccta ctagttcaat gactattaaa     420
acaggaaagc atgttgatgc aaatgggaag atatatcttc cttatctaca tgaatggaaa     480
cacccacagt cagacttgtt ggggcttatt caggtcatga ttgtggtatt tggagatgaa     540
cctccagtct tctctcgtcc tatttcggca tcctatccgc cataccaggc aacggggcca     600
ccaaatactt cctacatgcc aggcatgcca ggtggaatct ctccataccc atccggatac     660
cctcccaatc ccagtggtta cccaggctgt ccttacccac ctggtggtcc atatcctgcc     720
acaacaagtt ctcagtaccc ttctcagcct cctgtgacca ctgttggtcc cagtagggat     780
ggcacaatca gcgaggacac catccgagcc tctctcatct ctgcggtcag tgacaaactg     840
agatggcgga tgaaggagga aatggatcgt gcccaggcag agctcaatgc cttgaaacga     900
acagaagaag acctgaaaaa gggtcaccag aaactggaag agatggttac ccgtttagat     960
caagaagtag ccgaggttga taaaaacata gaacttttga aaaagaagga tgaagaactc    1020
agttctgctc tggaaaaaat ggaaaatcag tctgaaaaca atgatatcga tgaagttatc    1080
attcccacag ctcccttata caaacagatc ctgaatctgt atgcagaaga aaacgctatt    1140
gaagacacta tcttttactt gggagaagcc ttgagaaggg gcgtgataga cctggatgtc    1200
ttcctgaagc atgtacgtct tctgtcccgt aaacagttcc agctgagggc actaatgcaa    1260
aaagcaagaa agactgccgg tctcagtgac ctctactgac ttctctgata ccagctggag    1320
gttgagctct tcttaaagta ttcttctctt ccttttatca gtaggtgccc agaataagtt    1380
attgcagttt atcattcaag tgtaaaatat tttgaatcaa taatatattt tctgttttct    1440
```

tttggtaaag actggctttt attaatgcac tttctatcct ctgtaaactt tttgtgctga 1500
atgttgggac tgctaaataa aatttgttgc ataaaaaaaa aaaaaaaaaa 1550

<210> 3
<211> 390
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:/note =
synthetic construct

<400> 3
Met Ala Val Ser Glu Ser Gln Leu Lys Lys Met Val Ser Lys Tyr Lys
1 5 10 15
Tyr Arg Asp Leu Thr Val Arg Glu Thr Val Asn Val Ile Thr Leu Tyr
20 25 30
Lys Asp Leu Lys Pro Val Leu Asp Ser Tyr Val Phe Asn Asp Gly Ser
35 40 45
Ser Arg Glu Leu Met Asn Leu Thr Gly Thr Ile Pro Val Pro Tyr Arg
50 55 60
Gly Asn Thr Tyr Asn Ile Pro Ile Cys Leu Trp Leu Leu Asp Thr Tyr
65 70 75 80
Pro Tyr Asn Pro Pro Ile Cys Phe Val Lys Pro Thr Ser Ser Met Thr
85 90 95
Ile Lys Thr Gly Lys His Val Asp Ala Asn Gly Lys Ile Tyr Leu Pro
100 105 110
Tyr Leu His Glu Trp Lys His Pro Gln Ser Asp Leu Leu Gly Leu Ile
115 120 125
Gln Val Met Ile Val Val Phe Gly Asp Glu Pro Pro Val Phe Ser Arg
130 135 140
Pro Ile Ser Ala Ser Tyr Pro Pro Tyr Gln Ala Thr Gly Pro Pro Asn
145 150 155 160
Thr Ser Tyr Met Pro Gly Met Pro Gly Gly Ile Ser Pro Tyr Pro Ser
165 170 175
Gly Tyr Pro Pro Asn Pro Ser Gly Tyr Pro Gly Cys Pro Tyr Pro Pro
180 185 190
Gly Gly Pro Tyr Pro Ala Thr Thr Ser Ser Gln Tyr Pro Ser Gln Pro
195 200 205
Pro Val Thr Thr Val Gly Pro Ser Arg Asp Gly Thr Ile Ser Glu Asp
210 215 220
Thr Ile Arg Ala Ser Leu Ile Ser Ala Val Ser Asp Lys Leu Arg Trp
225 230 235 240
Arg Met Lys Glu Glu Met Asp Arg Ala Gln Ala Glu Leu Asn Ala Leu
245 250 255
Lys Arg Thr Glu Glu Asp Leu Lys Lys Gly His Gln Lys Leu Glu Glu
260 265 270
Met Val Thr Arg Leu Asp Gln Glu Val Ala Glu Val Asp Lys Asn Ile
275 280 285
Glu Leu Leu Lys Lys Lys Asp Glu Glu Leu Ser Ser Ala Leu Glu Lys
290 295 300
Met Glu Asn Gln Ser Glu Asn Asn Asp Ile Asp Glu Val Ile Ile Pro
305 310 315 320
Thr Ala Pro Leu Tyr Lys Gln Ile Leu Asn Leu Tyr Ala Glu Glu Asn
325 330 335
Ala Ile Glu Asp Thr Ile Phe Tyr Leu Gly Glu Ala Leu Arg Arg Gly
340 345 350
Val Ile Asp Leu Asp Val Phe Leu Lys His Val Arg Leu Leu Ser Arg
355 360 365

Lys Gln Phe Gln Leu Arg Ala Leu Met Gln Lys Ala Arg Lys Thr Ala
370 375 380
Gly Leu Ser Asp Leu Tyr
385 390

<210> 4
<211> 437
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:/note =
synthetic construct

<400> 4
Met Thr Thr Ser Thr Leu Gln Lys Ala Ile Asp Leu Val Thr Lys Ala
1 5 10 15
Thr Glu Glu Asp Lys Ala Lys Asn Tyr Glu Glu Ala Leu Arg Leu Tyr
20 25 30
Gln His Ala Val Glu Tyr Phe Leu His Ala Ile Lys Tyr Glu Ala His
35 40 45
Ser Asp Lys Ala Lys Glu Ser Ile Arg Ala Lys Cys Val Gln Tyr Leu
50 55 60
Asp Arg Ala Glu Lys Leu Lys Asp Tyr Leu Arg Ser Lys Glu Lys His
65 70 75 80
Gly Lys Lys Pro Val Lys Glu Asn Gln Ser Glu Gly Lys Gly Ser Asp
85 90 95
Ser Asp Ser Glu Gly Asp Asn Pro Glu Lys Lys Lys Leu Gln Glu Gln
100 105 110
Leu Met Gly Ala Val Val Met Glu Lys Pro Asn Ile Arg Trp Asn Asp
115 120 125
Val Ala Gly Leu Glu Gly Ala Lys Glu Ala Leu Lys Glu Ala Val Ile
130 135 140
Leu Pro Ile Lys Phe Pro His Leu Phe Thr Gly Lys Arg Thr Pro Trp
145 150 155 160
Arg Gly Ile Leu Leu Phe Gly Pro Pro Gly Thr Gly Lys Ser Tyr Leu
165 170 175
Ala Lys Ala Val Ala Thr Glu Ala Asn Asn Ser Thr Phe Phe Ser Val
180 185 190
Ser Ser Ser Asp Leu Met Ser Lys Trp Leu Gly Glu Ser Glu Lys Leu
195 200 205
Val Lys Asn Leu Phe Glu Leu Ala Arg Gln His Lys Pro Ser Ile Ile
210 215 220
Phe Ile Asp Glu Val Asp Ser Leu Cys Gly Ser Arg Asn Glu Asn Glu
225 230 235 240
Ser Glu Ala Ala Arg Arg Ile Lys Thr Glu Phe Leu Val Gln Met Gln
245 250 255
Gly Val Gly Asn Asn Asn Asp Gly Thr Leu Val Leu Gly Ala Thr Asn
260 265 270
Ile Pro Trp Val Leu Asp Ser Ala Ile Arg Arg Arg Phe Glu Lys Arg
275 280 285
Ile Tyr Ile Pro Leu Pro Glu Glu Ala Ala Arg Ala Gln Met Phe Arg
290 295 300
Leu His Leu Gly Ser Thr Pro His Asn Leu Thr Asp Ala Asn Ile His
305 310 315 320
Glu Leu Ala Arg Lys Thr Glu Gly Tyr Ser Gly Ala Asp Ile Ser Ile
325 330 335
Ile Val Arg Asp Ser Leu Met Gln Pro Val Arg Lys Val Gln Ser Ala
340 345 350

```
Thr His Phe Lys Lys Val Cys Gly Pro Ser Arg Thr Asn Pro Ser Met
        355                 360                 365
Met Ile Asp Asp Leu Leu Thr Pro Cys Ser Pro Gly Asp Pro Gly Ala
    370                 375                 380
Met Glu Met Thr Trp Met Asp Val Pro Gly Asp Lys Leu Leu Glu Pro
385                 390                 395                 400
Val Val Cys Met Ser Asp Met Leu Arg Ser Leu Ala Thr Thr Arg Pro
                405                 410                 415
Thr Val Asn Ala Asp Asp Leu Leu Lys Val Lys Lys Phe Ser Glu Asp
            420                 425                 430
Phe Gly Gln Glu Ser
        435
```

<210> 5
<211> 437
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:/note =
synthetic construct

<400> 5

```
Met Thr Thr Ser Thr Leu Gln Lys Ala Ile Asp Leu Val Thr Lys Ala
 1               5                  10                  15
Thr Glu Glu Asp Lys Ala Lys Asn Tyr Glu Glu Ala Leu Arg Leu Tyr
            20                  25                  30
Gln His Ala Val Glu Tyr Phe Leu His Ala Ile Lys Tyr Glu Ala His
        35                  40                  45
Ser Asp Lys Ala Lys Glu Ser Ile Arg Ala Lys Cys Val Gln Tyr Leu
    50                  55                  60
Asp Arg Ala Glu Lys Leu Lys Asp Tyr Leu Arg Ser Lys Glu Lys His
65                  70                  75                  80
Gly Lys Lys Pro Val Lys Glu Asn Gln Ser Glu Gly Lys Gly Ser Asp
                85                  90                  95
Ser Asp Ser Glu Gly Asp Asn Pro Glu Lys Lys Lys Leu Gln Glu Gln
            100                 105                 110
Leu Met Gly Ala Val Val Met Glu Lys Pro Asn Ile Arg Trp Asn Asp
        115                 120                 125
Val Ala Gly Leu Glu Gly Ala Lys Glu Ala Leu Lys Glu Ala Val Ile
    130                 135                 140
Leu Pro Ile Lys Phe Pro His Leu Phe Thr Gly Lys Arg Thr Pro Trp
145                 150                 155                 160
Arg Gly Ile Leu Leu Phe Gly Pro Pro Gly Thr Gly Lys Ser Tyr Leu
                165                 170                 175
Ala Lys Ala Val Ala Thr Glu Ala Asn Asn Ser Thr Phe Phe Ser Val
            180                 185                 190
Ser Ser Ser Asp Leu Met Ser Lys Trp Leu Gly Glu Ser Glu Lys Leu
        195                 200                 205
Val Lys Asn Leu Phe Glu Leu Ala Arg Gln His Lys Pro Ser Ile Ile
    210                 215                 220
Phe Ile Asp Glu Val Asp Ser Leu Cys Gly Ser Arg Asn Glu Asn Glu
225                 230                 235                 240
Ser Glu Ala Ala Arg Arg Ile Lys Thr Glu Phe Leu Val Gln Met Gln
                245                 250                 255
Gly Val Gly Asn Asn Asn Asp Gly Thr Leu Val Leu Gly Ala Thr Asn
            260                 265                 270
Ile Pro Trp Val Leu Asp Ser Ala Ile Arg Arg Arg Phe Glu Lys Arg
        275                 280                 285
```

```
Ile Tyr Ile Pro Leu Pro Glu Glu Ala Ala Arg Ala Gln Met Phe Arg
    290                 295                 300
Leu His Leu Gly Ser Thr Pro His Asn Leu Thr Asp Ala Asn Ile His
305                 310                 315                 320
Glu Leu Ala Arg Lys Thr Glu Gly Tyr Ser Gly Ala Asp Ile Ser Ile
                325                 330                 335
Ile Val Arg Asp Ser Leu Met Gln Pro Val Arg Lys Val Gln Ser Ala
            340                 345                 350
Thr His Phe Lys Lys Val Cys Gly Pro Ser Arg Thr Asn Pro Ser Met
        355                 360                 365
Met Ile Asp Asp Leu Leu Thr Pro Cys Ser Pro Gly Asp Pro Gly Ala
    370                 375                 380
Met Glu Met Thr Trp Met Asp Val Pro Gly Asp Lys Leu Leu Glu Pro
385                 390                 395                 400
Val Val Cys Met Ser Asp Met Leu Arg Ser Leu Ala Thr Thr Arg Pro
                405                 410                 415
Thr Val Asn Ala Asp Asp Leu Leu Lys Val Lys Lys Phe Ser Glu Asp
            420                 425                 430
Phe Gly Gln Glu Ser
        435
```

<210> 6
<211> 444
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:/note =
synthetic construct

<400> 6

```
Met Ser Ser Thr Ser Pro Asn Leu Gln Lys Ala Ile Asp Leu Ala Ser
 1                 5                  10                  15
Lys Ala Ala Gln Glu Asp Lys Ala Gly Asn Tyr Glu Glu Ala Leu Gln
            20                  25                  30
Leu Tyr Gln His Ala Val Gln Tyr Phe Leu His Val Val Lys Tyr Glu
        35                  40                  45
Ala Gln Gly Asp Lys Ala Lys Gln Ser Ile Arg Ala Lys Cys Thr Glu
    50                  55                  60
Tyr Leu Asp Arg Ala Glu Lys Leu Lys Glu Tyr Leu Lys Asn Lys Glu
65                  70                  75                  80
Lys Lys Ala Gln Lys Pro Val Lys Glu Gly Gln Pro Ser Pro Ala Asp
                85                  90                  95
Glu Lys Gly Asn Asp Ser Asp Gly Glu Gly Glu Ser Asp Asp Pro Glu
            100                 105                 110
Lys Lys Lys Leu Gln Asn Gln Leu Gln Gly Ala Ile Val Ile Glu Arg
        115                 120                 125
Pro Asn Val Lys Trp Ser Asp Val Ala Gly Leu Glu Gly Ala Lys Glu
    130                 135                 140
Ala Leu Lys Glu Ala Val Ile Leu Pro Ile Lys Phe Pro His Leu Phe
145                 150                 155                 160
Thr Gly Lys Arg Thr Pro Trp Arg Gly Ile Leu Leu Phe Gly Pro Pro
                165                 170                 175
Gly Thr Gly Lys Ser Tyr Leu Ala Lys Ala Val Ala Thr Glu Ala Asn
            180                 185                 190
Asn Ser Thr Phe Phe Ser Ile Ser Ser Ser Asp Leu Val Ser Lys Trp
        195                 200                 205
Leu Gly Glu Ser Glu Lys Leu Val Lys Asn Leu Phe Gln Leu Ala Arg
    210                 215                 220
```

Glu Asn Lys Pro Ser Ile Ile Phe Ile Asp Glu Ile Asp Ser Leu Cys
225 230 235 240
Gly Ser Arg Ser Glu Asn Glu Ser Glu Ala Ala Arg Arg Ile Lys Thr
245 250 255
Glu Phe Leu Val Gln Met Gln Gly Val Gly Val Asp Asn Asp Gly Ile
260 265 270
Leu Val Leu Gly Ala Thr Asn Ile Pro Trp Val Leu Asp Ser Ala Ile
275 280 285
Arg Arg Arg Phe Glu Lys Arg Ile Tyr Ile Pro Leu Pro Glu Pro His
290 295 300
Ala Arg Ala Ala Met Phe Lys Leu His Leu Gly Thr Thr Gln Asn Ser
305 310 315 320
Leu Thr Glu Ala Asp Phe Arg Glu Leu Gly Arg Lys Thr Asp Gly Tyr
325 330 335
Ser Gly Ala Asp Ile Ser Ile Ile Val Arg Asp Ala Leu Met Gln Pro
340 345 350
Val Arg Lys Val Gln Ser Ala Thr His Phe Lys Lys Val Arg Gly Pro
355 360 365
Ser Arg Ala Asp Pro Asn His Leu Val Asp Asp Leu Leu Thr Pro Cys
370 375 380
Ser Pro Gly Asp Pro Gly Ala Ile Glu Met Thr Trp Met Asp Val Pro
385 390 395 400
Gly Asp Lys Leu Leu Glu Pro Val Val Ser Met Ser Asp Met Leu Arg
405 410 415
Ser Leu Ser Asn Thr Lys Pro Thr Val Asn Glu His Asp Leu Leu Lys
420 425 430
Leu Lys Lys Phe Thr Glu Asp Phe Gly Gln Glu Gly
435 440

<210> 7
<211> 437
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:/note =
synthetic construct

<400> 7
Met Thr Thr Ser Thr Leu Gln Lys Ala Ile Asp Leu Val Thr Lys Ala
1 5 10 15
Thr Glu Glu Asp Lys Ala Lys Asn Tyr Glu Glu Ala Leu Arg Leu Tyr
20 25 30
Gln His Ala Val Glu Tyr Phe Leu His Ala Ile Lys Tyr Glu Ala His
35 40 45
Ser Asp Lys Ala Lys Glu Ser Ile Arg Ala Lys Cys Met Gln Tyr Leu
50 55 60
Asp Arg Ala Glu Lys Leu Lys Asp Tyr Leu Arg Asn Lys Glu Lys His
65 70 75 80
Gly Lys Lys Pro Val Lys Glu Asn Gln Ser Glu Gly Lys Gly Ser Asp
85 90 95
Ser Asp Ser Glu Gly Asp Asn Pro Glu Lys Lys Lys Leu Gln Glu Gln
100 105 110
Leu Met Gly Ala Val Val Met Glu Lys Pro Asn Ile Arg Trp Asn Asp
115 120 125
Val Ala Gly Leu Glu Gly Ala Lys Glu Ala Leu Lys Glu Ala Val Ile
130 135 140
Leu Pro Ile Lys Phe Pro His Leu Phe Thr Gly Lys Arg Thr Pro Trp
145 150 155 160

```
Arg Gly Ile Leu Leu Phe Gly Pro Pro Gly Thr Gly Lys Ser Tyr Leu
                165                 170                 175
Ala Lys Ala Val Ala Thr Glu Ala Asn Asn Ser Thr Phe Phe Ser Val
            180                 185                 190
Ser Ser Ser Asp Leu Met Ser Lys Trp Leu Gly Glu Ser Glu Lys Leu
        195                 200                 205
Val Lys Asn Leu Phe Glu Leu Ala Arg Gln His Lys Pro Ser Ile Ile
    210                 215                 220
Phe Ile Asp Glu Val Asp Ser Leu Cys Gly Ser Arg Asn Glu Asn Glu
225                 230                 235                 240
Ser Glu Ala Ala Arg Arg Ile Lys Thr Glu Phe Leu Val Gln Met Gln
                245                 250                 255
Gly Val Gly Asn Asn Asn Asp Gly Thr Leu Val Leu Gly Ala Thr Asn
            260                 265                 270
Ile Pro Trp Val Leu Asp Ser Ala Ile Arg Arg Arg Phe Glu Lys Arg
        275                 280                 285
Ile Tyr Ile Pro Leu Pro Glu Glu Ala Ala Arg Ala Gln Met Phe Arg
    290                 295                 300
Leu His Leu Gly Ser Thr Pro His Asn Leu Thr Asp Ala Asn Ile His
305                 310                 315                 320
Glu Leu Ala Arg Lys Thr Glu Gly Tyr Ser Gly Ala Asp Ile Ser Ile
                325                 330                 335
Ile Val Arg Asp Ser Leu Met Gln Pro Val Arg Lys Val Gln Ser Ala
            340                 345                 350
Thr His Phe Lys Lys Val Cys Gly Pro Ser Arg Thr Asn Pro Ser Val
        355                 360                 365
Met Ile Asp Asp Leu Leu Thr Pro Cys Ser Pro Gly Asp Pro Gly Ala
    370                 375                 380
Ile Glu Met Thr Trp Met Asp Val Pro Gly Asp Lys Leu Leu Glu Pro
385                 390                 395                 400
Val Val Cys Met Ser Asp Met Leu Arg Ser Leu Ala Thr Thr Arg Pro
                405                 410                 415
Thr Val Asn Ala Asp Asp Leu Leu Lys Val Lys Lys Phe Ser Glu Asp
            420                 425                 430
Phe Gly Gln Glu Ser
        435
```

<210> 8
<211> 37620
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:/note =
synthetic construct

<400> 8

```
ggatccaatc cagaatccca tactgcattt agttgtcatc ttcttagtct ctacaatctg      60
tgacagttcc tcagtctttc cttgtctttt atgatcttga cattttgag gagcactggt      120
tattttgtag aagtagaaaa accctccata tgaatttttc tgatgttttc tcatgattag      180
attgaggtta caggtttggg aagcatacca cagagctaat gtgcccttct ccacatcagg      240
ggttccatca taacaacatg gcttattact ggtgataaaa gaggggtttt ttgtttgttt      300
gtttgttttt accgattctc actctgttgc ccaggctgaa gtgcggtggt gtgatcttgg      360
ctcactgcaa tctctgcctc ctgggttcta gtgattctcc tgcttcaccc tcccaagtag      420
ctggagttac atgcaggcac catcacgcct ggctaatttt tgtttgtttg ttttgagaca      480
gagtttcgct ctgtagccca ggctggagtg ctgtggtgcg atctcagctc accacaacct      540
ctgcctccta ggtcccggtt caagaaattc tcctgcctca gcctcccgag tagctgggat      600
tacaggaacg caccaccatg cccagctaat ttttgtattt ttagtagaga cagggtttca      660
ccatgttggc caggcgggtc ttgaactgct gacctcaggt gatcgcccgc cttggcctcc      720
```

caaagtgctg ggattacagg cgtgactcac cacacccggc ctaatttttg tatttttagt 780
agagatggag tttcaccatg ttggccaggc tggtctcgaa ctcctgacct caagtgatcc 840
acccacctcg gcctcccaaa gtgctgggat tacaggcgtg agtcaccacg cccggcccat 900
ctttgcttaa attataaact cagtttctcc cacagttagt tcagcctatg ctcaggaatg 960
tacaaggaca gcttggaggt taaaagtaag atggagttgg ttaagttaga tctctttcac 1020
tgtctcagtc atcattttgc gaaggcaaat tcacaagtgg gaagaaagga cgggtagatg 1080
cagggcagag agagggctct gtgcttcgag cctggatgat acagatttgg caatactatt 1140
ctaaggagag accagagaaa gaagtttagg gtagaaggta attaattaga gtatgttgaa 1200
tttgtcttgc agaaaaaccc tgtgaaggaa acctctggga ggctgttgta aatgctgggc 1260
tggggatggg gtgggtctag agatccagat ttgaagcaat ttctagtcta ttgatagctg 1320
aagctcaaga agagagttag atgagcagag gtgaagagag aagaggacat cagactcttg 1380
agtaaattca gcttaagctc agggaggagt aagaaaggga gacagtgaag gaagcagata 1440
aggaatattg ggaagctgga caaacaagaa ctgggtagtt gacaatacca atgaagaggg 1500
aagaggagag gactttaatt ctttcggaga tgccagatga gaaaaaatta agaaatgagt 1560
gagaaaaaaa tgatggagtg attaaacatc tatgttgctt tcctttgaga atgttaacta 1620
tgaaaagaag aaaagggtgg tggtggctgc aaaccatagg aaggctgaga aagggtttta 1680
ctttggttct gcctctaatc cctgacgcat tcaacctcct caagccattt gcttgtaaac 1740
ataactggca taaataatat ctgatatgcc cattttacag gacaagtgct tccctctcag 1800
accacatact cgaggacagc ggccgtccac gccttattta cattcatctt cccaacacaa 1860
cagtcatttg gaagatgccg agtggcgtgg gaagaacaag cgggaacgag cgccgacccg 1920
cccgcggcct gactgctagc gcaccgctgt gcggccgcct caccgcgcac tccctcctga 1980
attattcacc accgagacga ggacggtgat cgcccccatt ccacagagga cacggagtga 2040
aggcgcagtc acaggcccaa agcccccagt cggcgggtct tggaaatgga cgacttttc 2100
actcaaaagc tacgaggcaa caaagcatgt ggatgaggag ggaagggagc cggggccaat 2160
aattcatctc aaagtaaaag acgtccccgg gggtggggaa caaacaaaca acaaaacccg 2220
gctctgtgct tccgcgcttg cgctctccta ggtctgccag gctgccgccc gcgcctaggc 2280
cagagtctca gtgcgcctgc gccgagcccg acttccgcca gggctcgatt ggcttccagg 2340
ctacttctcc cggcccaggg agggcctgcg ctgtggctag agaagggagt ctggtgactg 2400
ggtgcgcgtg ctgatgacga aatcggaagt gcccggagca gaggcgggac cagaacctag 2460
agggcctcgg gattgcggaa gtttggtggg gagggtcgga gctctggtgg agagagtgtt 2520
gtctaaaaca agttccggaa gggaggctgc ccttcgcggt ccgagaacca ccggcctccc 2580
cagtttgagg gctgttaccc cgtgcgcgct tcgacgttgc tgctgttggc tctcctcgcc 2640
cctcgttccc ttgggaaccg cctgggaact ccgccatgtc atccacttcg cccaacctcc 2700
aggtatcatt gctcatctcc cgtcccccta gggagctggg atcgagggga ggcgcgagaa 2760
gcgggtgggg ggcggggaga agggaggaag cggccgcgcc tgtggtcgag gcgcaggagg 2820
gcgggtccgg ggacctgcag caggtgggag aggggcccgt tttccaacgg aaaggtcgtt 2880
tctggcccgc tcttggggtc tcccgagcta ctggctttgt gagactcttc acttttgttg 2940
acacttcccg cttccaaacc agccaagatc tgggcaggtt tctccaatca gggttaacgg 3000
agaaagaaag aacgtaattt ttgcaaagcc ctattctgga ccagttgaag atagggtggt 3060
catagttcgg tcttagtttc tcaagctgac agcagacccc ttcattctga gctgtcaggt 3120
gagtagagat atgggaacag accaagaatc tgagattaca gccagcctgc atgcaatctc 3180
cgtgaagtca ttgaatttat gtgcacactc ctgtggctcc cagatggtag tgtgtaagat 3240
cgaggagtat tcttaaacct ctgtgtggct ccacttcccc tcctctaccg caatgagcac 3300
agtgttttgt acacaaatca tggaatctta agattttagg cttaatcgtt agtcctctgt 3360
ggaatggaat tcgttcagaa aagaaatttc ttaatgtaat ggtgtctccc atccccagaa 3420
tggatctagg gcttcagtgg cgtttagaat gtaatgtggc ttaatgtaga ggactcccag 3480
tccagtaagt caataaggca caaaagagta gcctaaaata cttgtattga gtccccgtcg 3540
tggacttccg gtagtttgtt gaacaagcag ttattggagt gcctgacaga ggcatcaaag 3600
gtaatggaag aactccttaa cccatgataa ttgattctga agaaagacca ctcaaagtag 3660
acccacttaa acagagatta aaatttcaca cacacataca caaaaagagt gggaaaggag 3720
gcttgcttgg gactttcatc ttgaaaatcg acctagaaga ttattaaaat aattgttatc 3780
tttcacaaaa tattataatt gtataattgc aacaatgact cattttgtgg aataagtcaa 3840
gtgttgattt gtctgtgttt ttttcctttt cttttctttc ttttttttt ttttttgag 3900
acggagtctt gcacggtcgc ccaggcttga gtgcagtggc gtgatttcag ctcactacaa 3960
cctccacctc ccaggttcaa gcgattctct tgcctcagcc tcctgagtac ctgggactac 4020
aggcacccac caccacgccc agctaatttt tgtattttta gtagtagaga tgaggtttca 4080
ccatgttggc ccaggctggc ctcttgacct caggtgatcc gcctacctca gcctcccaaa 4140
gtgctgggat tacaggcgtg agtcaccgtg cccggatgac tatctttttc accaaggttg 4200
tattaaacta agattcagca atcttcactt gagtttttg tagcttttg tagtcctatt 4260

tggaaaggtt ctttttggta aacgtatttg gaaaaagcgt tggtcatttg gtaaaagtat 4320
ttttcccttg agtattaaag gtgactcttc tgaagtgccc atcttgtcat cgttggcttt 4380
ttaacaattt ttattccacc agcttttcat ttgtcagaga ctttgcatgt gatttcgata 4440
ctctacaaca taatcttcat catttgcagg tttttgttgt tgttgttttg agacagtctt 4500
gctctgttgc ccagactgga gtgcagtggc gtgatcttgg cgcactgcaa cctccacctc 4560
ccaggttcaa gtgattcttg tgcctcagcc tcccgagtag ctgggattac aggtgcctgc 4620
caccacgcct ggctaatttt ttctattttt agtagagaca gggtttcacc atgttggcca 4680
cgctggtctt gaactcctga cctcaagtga tccgcccgcc ttagcctccc aaagtgctgg 4740
gttacaggtg tgagccaccg tgcccggccc cagaatgcat catacttgaa ttggactgtt 4800
ggctgttctt cagtgaatca gccagagact gaccaacatg aaaggcataa ttgttactct 4860
taagcaggaa ggatcttgaa ttgtaatttt gttagtggta ggttcacatg tgtgctttct 4920
cattcaacct tttcaacctt ctaacctagg aaatttggat taatattcat gtagtgatta 4980
acttcctgct gaacaaggca aaaatctctg tggccagatg cttataatct aatgagggaa 5040
tgcagataac taaaaatcaa aattatataa tgtaataaag ctgcacagta ttctccactt 5100
ccctgtggtt ttactttcca cagtttcaat aacccgaagt taacctgggt ccaaaaattt 5160
taaatgaaat attctagaat aaatagttca taattttgaa acactatgtc acaatgcctg 5220
tgtcattcac ttcagttcat atcatgacgt aggcatttta ttatttcaca taatcacaag 5280
aaggttgaga acagtgcaat aagatgttga gagagagaaa ccacattcac atcactttta 5340
ttacagtata ttgtaataaa atgtactggg caaggaggca aggtctgtgg gatgggaggc 5400
ggtaaggggc caaattagta gagtatccat tcaaatgaag tggaaaggat tgatcaaagg 5460
agaaaaattg aagatggaga agagtcaagc ttgtatctgc atttattttt tttttttga 5520
gacagagtct cactctattg cccaggcagg agtgcagtgg catgatctcg gcttatggta 5580
acctccatct cccaggttca agcgattccc ctgcatcagc cccccatgta gcagctggga 5640
ttacaggcat gcaccaacac acccagctaa tttttgtaga gatggggatt cgctgtgttg 5700
gccaggccag tctcaaagtc ctgaactcag gtgatccacc cccgttggcc tcccaaagtg 5760
gtagaattac aggcgtgagc cactgctccc tgctgtatct gcattttaaa agtgagggag 5820
gaggaagagt ctgagttttc tcccagattt ctgggttggg tatctgagta cccaactcag 5880
gtagggattg gttctgggaa tacaaaagga gagctatttg gggccagggc aagggagcaa 5940
aataatttgg ttctaaactt gtcgagtttg agatgcttgt ggagacaatc aaatggaaat 6000
gtccggtgaa tatttgcata catatgtgta gaagtaggga gagtggacta ggctggaaat 6060
aaggatttat tagcgtacag gtggtagcag aggttgtagg catacagagt cacatcctgg 6120
agagtttaca gaataaggca aaacaaagac cacgttagaa attctatgga ttagcgacag 6180
ttaaggagag agaaatggga gcctggaaaa gaattatgaa tcaaaagtaa aaccaggcca 6240
ggcgcagtgg ctcatgcctg taatcccagt gttttgggag gccgaggcag cagatcgctt 6300
cagcttagga cttcgagacc agtctgggca acatggtgaa aacccctctc tacaaaaatt 6360
agccaggtgt ggtggtgcat gcctgtagtc ccagctactc cagaggctga ggtgggagaa 6420
tggcttgagc ccaggaggtg gaggtaacac tgaggcgtga tcacacccgt gtactccagc 6480
ctgggtgaca cagtgagacc ctgtctaaaa aaaaaaaaat actaaaacta gaagaaaatg 6540
gtgtcatgac agccaggcaa ggagacactg tcagggtatt tgttgtgtgt gttaaatgct 6600
gcggagaggg agtgagataa gcatatgatt tggctatcag gaggcagaaa gcattgctgt 6660
gggagaaata ggagatcagg aaactgggat gaacatgggt attctttggg attttgctgt 6720
ggaggcaaag agatgagtga gatgtaggga gaggagaacc gaggagctga ggagtgatgg 6780
tgtgtagttt gtgttctgca gatgtgagtc tgagatagga caacacattt ttctgctgag 6840
ggaaaaaagg aaaggttgga gagaacagtt gataggcatg tcattgatgg agtaaatgtc 6900
agcagaggaa ggtggggatc agctccacag cagaggagaa agaattggtc tcaacagaca 6960
gggcctggga cacctcttct gaggtaggta attgtagaat ttaccccatt tctcagggtt 7020
ttctaaggtg gatgtgtctg gtactcctaa gtggaacagg acaattccac atgttcccaa 7080
cctttatact tccgtatgga taactggatc ctgactgatc ctggggtgac acagtcattc 7140
agtcttgtgt cagccactgt gctagactct gggaatataa agaagtgaag acaacagaca 7200
gaactgctct tctcctggag cttacattct aatagaagag agaggccaat aaattaaaaa 7260
aaaacaaaaa aaacatatgc atgctcagac acacacatgc acacacaaag atcatcagcc 7320
ttaagtcact gatgtagggg gggaggaaac agagagaaat gagtcccagc ccacacctga 7380
tgaatatact tcattgtcac ataagagagg aagaatcaat gaaggacgct gagaaggagg 7440
agccagaatg gtgggaggaa aaccagcaaa gtggtggttg ggggctcagc agaggacagt 7500
gtccaagaag gaggaaataa tcaactttgt aaaatgtcag aagatcaaga agacagagct 7560
ggtgtttgga tctagcatgc tagagagcct tgatatgagc cgttttatta tgtgatacag 7620
gaaagagaat agaggccact ctggtgagat ttttggctct aaaaggtagc agagaaacag 7680
aactaacacc gaacagtgtt ttgtcaaata agactttta aaggagggct gtttgtatgt 7740
cgatggaaat atctagtgga gagaaattgt taagaatgac agaagtgaag taaccaaaag 7800

gaagaagata ctgaaaaagg agtagggaat ttttttttc attcgttaaa caatttcgag 7860
tacctgttgt atgccggcac cattctaggc atgggtagct caaggtgaag tcctgtcctc 7920
atggaaccta cctccctaaa ggagacagcc agtgggtgag catacaggta aatacaatga 7980
cagtgatggg atgagagagt gcagtggggg cagatagcaa agagaaggaa tgaggtgcag 8040
cctttgctat agcaaattag ggaggcctct ctaaggactt gccatctgaa ctaggactga 8100
aaaggagaca tcatgcatct gtctggggaa aggagggaag agcattctag gtggaagggt 8160
tcaactgctt gggctgcctg agtgcaggag gattctaggg ttcggggaga accacacgta 8220
gctagggccc aggttatttc agaccttgta gatcagctct agcatcagga ttttattgtt 8280
tttaatggga gaccatcaag gttacctttt atgttttgag gtagcaaggg tgggaacaga 8340
ctgaccactc acgaggctgg gtagtagttt tgacatatat gatagcagct tgaactaggg 8400
tgttagcaag ggagatggtg aggattatat tcagaaagaa aattagctct ggctgaaaag 8460
ggaagataat gtttgatgga ccagagcaga tggagctgga agaggggaag acaggcaaca 8520
gactctgcca gtattatctg tctgagactg aagaactcca gtcagtagtg aatgataaca 8580
gccaggacct ccagaggctg caggcataaa tgaatgaact caatgcaaaa gttcacctat 8640
tgcaggagga gctacagctg ctgcaggaac agggctcctg tataagggaa gtagtccggg 8700
ccatgaataa ggagaaagtg ttagtcaagg tgtatcccaa gggcaagttt gtcttagatg 8760
tggacagaaa cagcaacatc aatgacatga gacccaattg ccaggggggc gctaagaaat 8820
gatagcgaca cactgcacaa tatcataccc agcaaagtag acccactggt gtcaccagtg 8880
atggtggaga aggtgccaac cttgatttaa gagatgattg gtggactgga caagcagatt 8940
aaggacatca cagaagtgat cgagcagcct gttaagcatt ccaagctctt tgaagcactg 9000
ggtatcacac aacccagggt ggtactgctg tacggacctc taggcactgg gaagacactg 9060
tttgcctagg ctgtggctca ccatacagac tgtacctttg ttcatatccc tggctctgaa 9120
tttgtataga aattcatcgg ggaagggcca agaatggtga gggagctgtt tgtcatggca 9180
tgaggacctg ccatctgtca tcagctccat cagctcctcg cagctggagg ggggttctgg 9240
aggggacagt gaagtgctgc acatcgtgct ggaactgctc aaccaactgg acagctttga 9300
ggcccccaag aagatcaagg ttatcatgga ccactgcttc actcagggca catcaacaga 9360
aaaattgaat tccagctgga cattttgaag atccattctc tgaaaaggaa cctgacccag 9420
gggatcaacc agagaaaaat tgccaagctt atgccaggag catcaggggc tgaagtgaaa 9480
gtcatataca gagaaagcag catatatgcc ctgagggaac ggcgagtgca tgtcacccag 9540
gaggactttg agatagcagt agccagcgtc atgcagaagg atagagagta aaacatgtcc 9600
atcaagaaac tatggaagtg cggggatgtc ctttgtatgg atccctccgg taaagctttc 9660
tggaatgaaa aaataaaata attagctctc cagaattaac catagaagac atggttaatt 9720
ttatttcaaa tgaaaaatgc taacagacat ggttaatttt atttcagatg aaaaatgtta 9780
acagaaataa ggttgaaatg gtattcatca gcagaaaaaa aacagatgca tatcctctta 9840
attcatcgtc ttcgattaaa ggtttaacgt gtgccaccca gtacagtggc tgcatgtggc 9900
tattttaatt acataaactc actcctcagt cacactagcc acatttcaag tgctcagaag 9960
ccatatacag acattctcat catcagggaa aggctggttt aattagcact atgtttggag 10020
ctagatgctt aaataatacc cactgcagta taagagattt aacttaataa aaatcagtgt 10080
ggtgcctcag tgaggcctga ctgaatgaga ggaacaagtg gctgaagagg tgagtggcac 10140
agcatttcaa aatcacctga gaagaaaatt atagttagga taagaaagag cttatttcta 10200
ggccaggcac ggtggctcac gcctgttatc ccagcacttt gggaggctga ggcaggtgta 10260
tcacctgagg tcaggagttt gagaccagcc tggccaacat ggtgaaatcc catctctact 10320
aaaaatacaa aaattagcca ggcatggtgg tgggtgcctg taatcccagc cactggggag 10380
gctgaggcag gagaatcaca tgaacctggg aggcagagat tgcagtgagc caagatggcg 10440
ccattgcact ccagtctggg cgacaagagt gagactccat ctcaaaagaa aaaagaaaaa 10500
ggatttattt ctgctatgtg gcttaaccac tgaggctttg agatgcattt ttttttttct 10560
tttttgagac agagtcttgc tctgtcaccc aggctagagt gcagtggcat gatctcggct 10620
cactgcaacc tctgtctccc ggattcaagc gattctcctg actcagcctc ccaagtagct 10680
gggattacag gcgcctggct aatatttttt gtatttttag agatggggtt tcattatatt 10740
ggccaggctg gtctcgaact cctgacctca agtgatccgc cctcctcggc ctcccatact 10800
gctgggatta caggcgtgag ccactgtgct tggcctcact ctgtcaccca ggctggagtg 10860
cagtggcaca atcttggctg tggtgcattt ttatatgacc agtctgacct aaccaccact 10920
tgactgtgtg actccatgtt ttatgttcat atactagctg ataaaacttg aaaatttagc 10980
tttgaattta ggctacttga cctgtgagat acgtaactat tacattttat caattctgaa 11040
gtaaaccatt tatttttctt ccctgctaca atagaggaag catggtcttc cagcgggaga 11100
cctgttcctc ccatttacag atctcagccc agttatatcc aatccatgct caactgcccc 11160
aaaatgtaat attcttacac ataattggtt tttgcattgt ttgaatgatt ttttgcttat 11220
gttatatact tacttttaa tatctctaac aactgtcaat tttcattttt ttttactggc 11280
acctatttct gttgaatagt tttactacct aaattaaccc agttggaatg ttttaataag 11340

tcatgtttgc cctttatttg tatgtatctg tgcataattt ttttttttt ttttttgagg 11400
tggagtctca ctctgttgcc caggctggag tgcagtggca tgatcttggc tcactgcaac 11460
caccgcctcc cgggttcaag caattctcct gcctcagtct cccaagtagc tgggactata 11520
ggcgcatgcc gctacgcctg gtcaattttt tgtgttttag tagagacggg gtttcactgt 11580
gttgcccacg ctggtcgcaa actcctgagc tcaggagaat tttagtttga gataattgta 11640
gattcacacg cagttgtaag aaatactgag aatccagcat tccctttatc cagttgtctg 11700
cagtggcagc atcttgtaga actgggacaa tatcacaacc aggatattga tattgataca 11760
ttccaccact cttactcata ttttcccagt tttacttgtc cttatgtgtt tttatgtgtg 11820
tgggtgtggt tgtttgtgca gttttatcac atgtagatat gtgttacatt ttgagagttc 11880
tttatgtatt ctcgatacta gttctttgtt agatatgttt tgcaaatatt tctcccagtc 11940
tatagattgt cctttcatcc ttttaacaaa atcttccaga gagcaaaagt tttgtcataa 12000
gatccagtta cctttttttt tcttttatat agtatgcttt tgatgtcaaa tctaagaact 12060
ctgcgcatcc ccagatccta agttttccat gttttttcta aaagttttag ctgggtgcgg 12120
tggtgtggac ttgtgatcct gctatttggg gaggctgatg caggatgatc acctgagccc 12180
aggagtttga ggctagcttg gtcgacatag tgagatggag ccccgtctca acaaataaat 12240
aagaatcaaa gttttgtagt tttacattta attctatgat ccatttttga gttaattttt 12300
gtaggaagta ttaagtttaa gtcagaattc ttttaggtgt atgcttatgg atgtctagtt 12360
cctctagcac cacttcttaa aaagcctgac tcggtgccat tgactggctt tttgcatcat 12420
tttcaaaaat ctcttggctt tattagtatg gatttatttc tgggactttt taaaaaaatg 12480
ccacaaaagt agtccatgtt ccttatggac atttttttaga agatgtgaaa aagtagaaaa 12540
tgaaaaatat atgttttatc aaagataaac tctgtttaac aatttccttc ccatcttttt 12600
cctgtgctgc ttttttctgc tgttgtttca aataaattcc tgctacatat aaaaagttgt 12660
gtcctgcttt ttaaaaatgc ccattattaa cactttccat attattacaa atatttctta 12720
atgttatttt agtgatcctt tggtgttctg ttttgtggca tcctttatga tttaagtttc 12780
caggttttgt ttttgttatt taaaagacct ctctgataat atctgtgctc agtaaacctt 12840
ttagttaatt actggattga aatatatgag tatttttcag attcttgata tttcattgcc 12900
ccattgcttc tgaagaggtt tcaccccatt tacatttcca ctattgttat atcttttaga 12960
ttaaagcagc atttaatttt tgagaatttc tctaaatgct tgtttagttt gaattttcac 13020
attctattat ttcttattag ccctaaattt aactttagct aagtgaaaat tacatgtaaa 13080
tgttagcgat tggtactaaa gtaacattca gtagcattta ttcctataac atgtaatttg 13140
catcttttca acttgtaatt tataataaaa tacagttgtg agttagaaaa aaatgatttt 13200
ttagttactg tttatttttt acttattttt gttttctttc acagcccagc aagctgagaa 13260
gctttaaaaa aaaaaaaaac tttcagtatg attatttact tcaaatttat gtttatgctt 13320
tgatttaaat gttgttattt tgcctccctc aaacagaaag cgatagatct ggctagcaaa 13380
gcagcgcaag aagacaaggc tgggaactac gaagaagccc ttcagctcta tcagcatgct 13440
gtgcagtatt ttcttcatgt cgttaaatgt gaggccatag gttatatttt atttaaaaac 13500
acgaaaagga aaatttattc tgtttcattt ctaaaaaatt gaaatataat gccagacaat 13560
tctaaacgat ccatttgtta aaaaaaaaaa atactccact ttctttctag tgatttttaa 13620
gtattgagat gtagctagca aatgtgacag ccatcatagt tattactttg tatatgtaaa 13680
catcttaaaa tgtagtctag tcagtggaaa ataggttagg agacattcta tttgtgtctt 13740
tacagatatg tattgcgtgt gtaccacgtg gcaggcactg gggatacaat ggaaagaaat 13800
ggttaagacc cctgccctca tagctcttaa atggaatgca gttagatgat tgaagatgtt 13860
cttttgagct tctcccattt agttctgagg ttttatgtta aagtgagact aagaatgagt 13920
cataacctgt aagctatgtg gggctcatcc catctattgc acgtgttgcc aaagttcact 13980
ttgaaaacag ctctgacaga ggcctttgtg tagtaaaatt tacgatgttg gtattacaca 14040
gaatatatgg gatttccatc actttactct tctgaaataa agcttttgta gacattcata 14100
gaaatgatta atactgcttt ccttctattt gtctttttta taattgggtt tgtctaagac 14160
aagcaacttt gaactatctg caaaaatgta agggaaaaat atgaaagctt cagaatatta 14220
aataatttcc tgttagttaa cattaactta atataccctg ctgtcttgct taggttgttt 14280
ctgattcaga cccactgttg gtctccaacc ctgctggagg aaggggtaga aaagaggtga 14340
taatcattgg ctatgatgag ttagagatga tcttccctcc ttcttccaac atagaatttc 14400
cttgtctgaa aaaatacata tttaagagtt caaacctttc tacatagaat actaatggaa 14460
aaaaatctct ctcatttccc tttaatagat gaagcacagg gtgataaagc caagcaaagt 14520
atcagggcaa agtgtacaga atatcttgat agagcagaaa aactaaagga gtacctgaaa 14580
aataaagaga aaaaagcaca gaagccagtg aaagaaggac agccgagtcc agcagatgag 14640
aaggggtacg tgtttaaaat tgttcaattt tattgtggct ttttgctttc gatttgtgaa 14700
atttcttttg attaaaaacag gtgtcagcaa actatggctc atgggccaaa gctgcccagc 14760
tgcctgtttt tggaataagc tttatcagaa cacagccgtg cccattcttt tacatattgt 14820
tgtctagggc tgctttcgtg ccgtaacaac aaagttgtgt ggttgtgtgg agaccgtatg 14880

gcccacgaaa cctaaaatat ttacttgttt ggtcctttac aaaaaaaagt ttgctgaccc 14940
ctagtttgcc aagttgaatg gataatattc ctcaaacaac ccttcctgac tgcctgctag 15000
taggaggaga gaatgcttgt aggaatagta ctcagttact ctctggccat tagctaatat 15060
ggaagcttaa ggttatagga gtgaccagag agttaaacct ttgaccacat ggaaaaaatg 15120
aaaagtaact ccaacaatgt aaatttttt tggatatctt atgcaaaatt ttttgatgtg 15180
aatttgtgtt agatttgctt tatgttgtca ttgtaaagaa aacaacacag cataacttat 15240
aaacgttgaa actgattttt ggcatgtctt cctagatggg ggatgtatca cttctagatg 15300
cagaatttgt ttattcagta catgacatat gcgtgccaag tttgtgctca gcactgggaa 15360
catattctaa ccacatcctc tctggctgtc ctattaattg atgccagtag ttccaggtaa 15420
cacaccctta agaaaatgag actctgggat ggagttatct ggcctgcttg ggtttgaaaa 15480
cagcgagatt ctgctcagct ttagaggatg tggttctcat tgttcatggt atgcagtggc 15540
gcaggttatt tagacaatac cctgtggtca caaggaatga agtgcctttt aggcagtgct 15600
ttggggcttc gtggctgcta ccacagacca tttggaagta catgaggaat ataagaacca 15660
gtgttgactt ggctgcttca aactatgtgg cagagtttga aggaaatgat aagcttaagc 15720
tcaaggtttt atattctctg ttcaagatgt gatcggggaa ccagggagtt tctgtgacac 15780
cctaagaaca tgttatcttc ttcatcgata gcactgactt aacaaaaatc agatgatagt 15840
ctgatcttac aaattactga attccaactt agggtttctt atatgaaagt gaggacattt 15900
ggtataaaac agggactcca aacatgaagt gaaaatattg ggaagaattt agttgacctt 15960
gactgcccag gacctccagc tgtcactgag actcccttga cagcagaagc ttctcttccc 16020
ttgtccatag aagctcttct ctcttgcttg aaaaccctgt aaggatctct tctgagatga 16080
ttacctgaca aggcgatgcc agtttttctt ttgtactgaa cagcaatctt cgtttacagt 16140
aggacccact ttggctaatt taagcaaaaa gggatatatt atagggtatt accttacaga 16200
atagtggggc cagggaatca ggagaaatgg ccaacagtac cagggttgtt caagtaggga 16260
tcccactgtt gccaactcct ctgctaagca ctgatggtgc ttgggagcag gtagtaaatc 16320
ctctgcgcag ctccctcagc aaaaccagtg ccttcgcctg aagattctgt ctccttgtga 16380
gcagccccca ccttcttgga agaatttagg tctcatgtgg aaccctagct acaagggaga 16440
tgagaaacgg cagggttttt tttgtttgtt tgtttttaat aaggacagcc tttatagtac 16500
tggacagcat gatagaaaga ggttggagtg agtgtcatgt gaatcttttg cagtgtctac 16560
ctcacccaca gccccaccgt cctatgataa ccaaggcaat tccattttgg cataaaacat 16620
gtcatgcttt tatctccctt tatatgttat tgttttaaaa atgatataat gccttatttt 16680
tggctttatc acttttatac ttaacattgt atcaggaata tctttccaat aaactatttc 16740
tacatcacgt ttaaatgttg caatgtatcc tgtcatatgc tagtccctat aatatgagtt 16800
acttttccc cttttgcaaa cagtgtgggt gtgattatcc ttatggtcaa atgcttgaac 16860
ttgtctaaaa ttatttcctc aagataaatc cctagatatg aatgaaatgc cttgatcgac 16920
ctacttgttc tttgggtgta gttcattgtc atcaggaagc ttcctctgtc ctttggaatc 16980
cccagactaa gggttgtaga cctttgtacc ttagtgttcc agatttttt cctcctaata 17040
tgacctcttg ttctaattta tatcttgccc agtcctgatg tttttactgt tgttgtttgc 17100
agctctcctg tgattcatct tgaatccttc atggagtgat gtgtggggca tgcatttatg 17160
aattaataca tacctaaata tgtcatggga atttggagct agttcctgac ttcattctgg 17220
tctttttgtc cctttctttt acatgattct ctattgatga taacgtgact gcttcttatt 17280
tactacattc tttaaaataa gtaagcataa cttttaatgt aaatgtctga ttttgccact 17340
gaaattacaa atatttcagg ttaatttttc cttcattttt gccttatttg atagcaatga 17400
gtgatcattt gaatatatta aattcttttt attaggaatg acagtgatgg ggaaggagaa 17460
tctgatgatc ctgaaaaaaa gaaactacag aatcaacttc aaggttgctt tagatttcat 17520
ttttaaattt actatcctat taaaaagact caaagtgaat tgtctatttg aattgttttt 17580
gttctccctc tatatttaaa tgaagtttaa gaactggctt tggttcccaa tagtcatgtg 17640
ctggctttgt ttagagtgtt ttcttagggt ttgccctcta ggatcatggt ttggtctttt 17700
ttatgtaata gttttgttgg cttcacatct gcctttaaaa ggctctcagt caacttgctc 17760
tttctttttg tttatttgcc gccttttaga gaaggtaaca cttatactat tagcacactg 17820
caaaatttga tggtattaaa atttactact gtgggagaag ttaggactta cacgtggctt 17880
ttattgccgt tgctgtagct cacttattga tggtttattt tgtttcttaa ctgcttcttca 17940
tctgcttttc agtgttcctg tttacattta attgtttaag ctttctgtgt ttaccatcca 18000
cacctgtaaa gccctttcat tctgttctac tatagcacag tggcataaca gaggggacat 18060
aggtttttgg ctccattcag atcttgattc tgttcaaggg tctgccactt agaactttct 18120
aattgtgggc aagtgactga atctctcaga accagcgtcc cccatctcta aaatagaaat 18180
gagcaatagc ttcctttcag ggcttttaag attaatgaca gtgtaaaaac tacctcatat 18240
atctgacgaa gtctagggac tcaataaatg gtagcactgt ttatctaaaa aggcatctat 18300
ttgttggctg gctttgttaa tcactgagat ttgattattc ttcgaaataa attatctaac 18360
aatttttcag attgaatttt ttgctcaccc attggtagta tgtattttgt tcatttgaat 18420

cttaccaaat agaatgcata actgggaagg atttccattt tataggttca gccttggtta 18480
ctgaatagtg aaaatactta cctgagccaa acttttagcc ttgctatact agtgtcttcg 18540
ttttggtagt atttttgtct tgcctaccct actgggatgt ctaggaaaac actgtaagga 18600
taatgtgttg acaacagacc aggtaatgaa tagagaacat ctttagatgc caagaacttc 18660
agtggtagat atagtccaag ggcaaaatct tgagtttaca ctcaagcgta gtggtgtccc 18720
caccacctga aactactggc acacactgcg cactcagttg gcatttattg ggtaacccag 18780
tgggtctgtg ccaagtggcc ttattccagt tgtaaagcaa gacacttttt ttgtctgctc 18840
tattggattt ggttttgttt ccctctccct accctgcccc aggtaaaaat cagttgcata 18900
tctttaattt tcgaaaagat acattttttt tttctttttt gttttttga gacaagggtg 18960
tcgctgtgtt gcccaggctg gagtgcagtg gcgcaatcac ggcttactgc agcctcgaac 19020
tcctgggttc cagggattct cctgcctcag cctcctgagt agctgggaac acaggcacac 19080
accaccatgc ccggctaagt ttttttttt gtttgtttgt ttgtttgttt gtttgttttt 19140
gtagaggtgg ggtttggcca tgttgcccag gctgctctcg aactcctgag ctcaagtgat 19200
ttgcccactt cagcctccca aagagctggg attacaattg tgagccaccg cacttggccg 19260
atacaaccag tttttaatag ggatttctct aaattcccta tcacttacaa taaaggccag 19320
aaatgtaaaa caaacaaata aaagagaaaa aaaaccaaat acttaaaaca tttaaatgag 19380
gaaatgtatt tgaaatttca ggatgaaata gaatatgaaa ggagtcctag ccttttatat 19440
ttgaatatcc tttttttttct actttttaat ggtttgctgc tggatttcat actattgtat 19500
ctttttttgaa tttttccttg ttttttgaac agtattttaa atcttttact tctgttttag 19560
attatgatat taagtgtata tttgttaaaa ttaaaaatcg gtttgtgatt tgagttttct 19620
gttagctaat tttattcagt gaatttttttg tgattaagta atatactagt aaatagtata 19680
gtcaaagacc ggataaattt atatcattgg aagtacattc catttgttag cattttttcat 19740
aatcccatcc ccccttttta attttgcaaa ccatgtctag tatgagattg tttatcagta 19800
aacattccaa agtcaccttt taacttagaa gtattatata tgaaaacaaa tcatgacaca 19860
tttttaaaat ttatagcaga aatcttgagt cagacggtat acatatataa aataaagttt 19920
atttttgttt aactgacatg aaatcctagg attactaatc ttaatgaaac ataaaaaaga 19980
cgaaatgaat ttacttttttt cttactttat agtagtctga acttttggaa tacaaaatgc 20040
ttatggtttc atgtaatttt aatctagtag ctaaagattt tcataattta gggaagtaag 20100
ttgctagtaa ctagtcttag agtaacagat tatctttcct aacaagaaga aacaatcagt 20160
agcagtgatt tcataaagta atagattgta taagtggtgg tgttctgtta cacttaaaac 20220
aactaactac atgtcatctt agaacaggct gcttcatagt ttatagtgga acagtttcca 20280
tattctagaa aattagaaaa gcttttgata aaagagaact agaaaatttg tcagtaggct 20340
atattaaaag acatccttct tggactctta ctgcaataat tgctttggct tttgaatata 20400
ctgcattcag aagctaaatc ttccctgtaa gttaactttg ttaattaaca aaagttgtaa 20460
atagatctat ttagcagcca aatagaaact tgggcaaatc ttaggaaagt attgcaaatc 20520
tgcaggggta tcttttctcg gtttcccatc atctgcttcc attccttcct ttcctctctc 20580
cttttccttc tcccttctgt ctgtctgcct atcttgagtt aaatttgtta agcatatttt 20640
ttttttcagt ggttgaggag aggattagat caccttttct tttggttaac aggagaaatg 20700
gtagaacaag tacttttgat actccaaaat caatgtattc aaatataacc accttatgga 20760
acgtttggtg cttttccaaa cactctaata ttcattatgt catttgattt tcgcaccaac 20820
ctatctaagt agtttgggca gatggtatta ttttcatttc acaagtgagg aagctgaaac 20880
agaagggtaa agtaacatgc cccaggttgt ctagtcatcc acctggttcc tagtcctatt 20940
gactactccc taacacctgc ttcctgctca cgctatttca aaaaggttgt tagtacataa 21000
aactttccca attttataat gcgtttcttt aacatcatta tttcatgttg tcttctaact 21060
ttggtgaaat agtctttctt aatttaaaga taacgtaatt ttttgcaggt gccattgtta 21120
tagaacgacc aaatgtgaaa tggagtgacg ttgctggact tgaaggagcc aaagaagcac 21180
tgaaagaggc tgtgatactg cctattaaat ttcctcatct ttttacaggt gacataactt 21240
ttaaataatt ttatttcatg agtttgtaaa agttcattga tgactgaggt gaagcattgt 21300
cattcttgtc actaatgaat tatcttaaat aaattgtgac tgttaaaagg tggtgatata 21360
tctttataca agatgtttta tttcttatac atatccatat ttgcagtgta tatatattgc 21420
aaaacaacct aaatgtcaac cagtagggta attcaccaac atgcagtgga atactatgca 21480
ctcagtaaaa tagagtgtac atctacatta ataaattgaa atctgcattg tattttattg 21540
aagaatgtcc acatctattg caaagatcaa aaagcagtta taaactgaat agtatgattc 21600
catgtttgct aaatgtgtat gctgagaagt aagtatagtt agattcacat agaaatatta 21660
atatatgcct aggggataag attataggag atttttacttt ttgactttca tctgaatttt 21720
tgcaatggtt atgattttta aagagtattt cttaaaaaga ggagtaagaa tcacttaaca 21780
gtttcagagg aagtgcataa cctctagtag caaaaggaca tttcattatc actgtgaaga 21840
gaatattgct gacctaccat ctggggaaac ttacaagttc tagttccagc agtaagtgat 21900
atgtctgtgc taaaccattc ctttatatta tccttcttgt gactgatagc cttgactgta 21960

atatggccac acccacagtc tgacagtcag aggtttttcc ttttgcaaac atgttggttt 22020
atgtcaacag caaatctgca atgttggcct cagactgaag agtccagaca gtcgtatcgt 22080
atagagcagt gatgaagagc aaaaagcctg taggccacgc actgtccgcc tagcttcaag 22140
gtgcagttta ctacccaact tcaggagtat gaaacttaac agcataaaag ttgacatcat 22200
aatcctataa aagcacaaaa atatggttga actcatttta tggtatatta cttataccac 22260
tttttctcaa acatgagtca ttgatgtatc atcttcatga atttttccat gtatgtgtac 22320
cattactaat ttttaatatt tttctttgct tatttttttag aacaacattt ttactagagt 22380
atagcataaa tatacagaga catacaggag cccctgggtg ttgggcttcc ttaattgtca 22440
caaagcaagc atgccagtgt catcaaccag atccagaaat aacagcattg ccagggacct 22500
ggcaggagcc ttctgcctca ttcctcccga gaggtcatct ctctctctcg acttctaacg 22560
ccagagatta ttttgcctgt atttgaactt taaatcatct cacatatttt tacttagtaa 22620
cgttattgaa aagactgctt tgtgccttta taagtggaaa aaaaagtgtt tttccaatac 22680
acattaaaaa attatttaaa gtatatacca cctaaaagta cttcatgtta taccagtggt 22740
atatgtacca tactcagaca ctgacttgta aatgtataca tgtaaaatag gaaaatgaaa 22800
ttcaagtaac agttgctttc ataacagtta agtgaaacag tgttattcat ctgcatatca 22860
ttcttttgca gatactgtgt acacatccat agtttgcacc tagcagtttt ttttgttttg 22920
ttttgttttg tttctttttt tgagacagag tttcactctt gtcacccagg ctggaatgca 22980
atggcgtggt ctcgtttcac tgcaacctcc acctcccgag ttcaagcaat tctcctgcct 23040
cagcctcccg agtatttggg attacaggca tctgccacca cacccagcta atttttgtat 23100
ttttactaga gtcagggttt caccatgttg gccacagttg tcccgagctc ctgacctcag 23160
gtgatccgcc catctcagcc tctcaaagtg ctgggattac atacatgagc caccatgccc 23220
agccatagtt ttttaaatgt atgtttgata gtttactgcc agctgtaacc cacctgtttt 23280
attttaggca ggattataat tgctcattag caaaattgag gctcgaagag ttgatgtatt 23340
cactggttat attttggcaa aactaggatt taaagtagta tcacctttt gttacccaac 23400
catgaaagtc tgctacctag ctggaggctg aggtgggcag atcacctgag gtcagaagat 23460
cgagaccagc ctgaccaaca tggtaaaacc ccgtctctac taaaatacaa aaactagcca 23520
ggcgtaatgg cgggcaccta taatctcagc tacctgggag gctgaggcgg gagaattgct 23580
tgaaaccagg aggcagaggt tgcagtgagc cgagatcgca ccactgccct ccagcctggg 23640
caacaagagt gagactctgt ctcaaaaaaa ataaaataaa agaaagtctg ctacctataa 23700
cactaacctg tgctttcctg tggaatcaag aattcagcag aaactcacaa gtcatccaaa 23760
ctttgttagg ttcacttttc caaagagcat aaatattcaa agaaaatctc ttccagctgc 23820
ccatttttta tggatataat ttttgtacaa tattctggta tcaagcagtg tacacagtag 23880
atgatcagaa agtacctttt aactgaactc ttttccccta ctagactcca cgcagggcag 23940
aaatcctttt gcctgtgcac atgttacggt tttgatacat attggttgaa tgaatgagca 24000
aacaaatgta cttcctgtta gaaggcttat gttatgcaag tctcttttttg ataaagtagt 24060
ctctagaaga aacaaaatat atttaagaca taccttagat tgatttattt ttaaaatatt 24120
gttttgatcc ttagaggttt agagtgaaaa tctttccaga gtattcctca gtatagatga 24180
tacaattaag tgttaaaatt ctcatgacat ttttctaaat ctaaatttta ggcaagagaa 24240
caccttggag gggaatccta ttatttgggc cgcctggaac aggaaagtcc tacttagcca 24300
aagctgtagc aacagaagcc aacaactcaa catttttttc aatatcttcc tctgatcttg 24360
tttctaagtg gctaggtgaa agtgaaaagt aagtagtaaa tcaatttttt aaagtttttt 24420
ctttgcctgt aattactgta tcactcctta atagatttct tacatcactt atagatattc 24480
agagaagtca tgtaatacag gtactcactt ctctgtcaaa tcagtctcta atctagaatt 24540
tcttttgtgg gaaaagatag catatatcta caacctagtg aatcataaag tgaattatcc 24600
agttgtagca ttcttatttt ctgtacctga aacatgatta ggctttttaa tctacagagt 24660
tcttgtaata ttgatatatt actttgtttt aatcttttac tatttttttag agacttaaaa 24720
gttttcattt atttttagac tggttaagaa tttattccaa cttgccagag agaacaagcc 24780
ctccattatc ttcattgatg aaattgattc tctctgtggt tcaagaagtg aaaatgaaag 24840
tgaagccgca cgtagaatta agacggagtt cctagtgcaa atgcaaggta gtgttactaa 24900
ttttttttt cttttgaga cggagtctca ctctgtcgcc caggctggag tgcagtggcg 24960
cgatctcagc tcactacaac ctccacctcc caggttcaag caattctcat gcctcagcct 25020
cctgagtagc tgggattaca ggcgcgcacc accatgcccg gctaatttaa tgttactgat 25080
ttttaaagtc gtttttgttt attttataac aatttttata acagtaatga aaaattcaag 25140
cagaatttct ccattctaac aaactgtttc attttgccat gtttcattcc aattgtcttc 25200
catatgcata cataattttt actcaattgt aattatcttg taggactcac tgacataaat 25260
agaaaagaga tttaatgtat tcatttataa tgtaccagga acaacttcag taaaagtaat 25320
atgtagactt cgaatgaaat ttttcaagaa atattttgaa acgaatgtta atagtatttt 25380
ccacaatctg ggagatttga ttcatatatt tgcatcatat tgcatcatag tgtatttgga 25440
tcatattgcc gtaaggtgat atttgattaa atacagttct tttccaaaga cagactctca 25500

agatgaagca ctttaatggt aaagcaccac atttcttaac agaatatatg gaagaagttt 25560
aaagaacacc aaaatgacaa attaattaaa gcattaccta ttttaatgca ggggttggtg 25620
tagacaatga tggaattttg gttctgggag ctacaaatat accctgggtt ctggattctg 25680
ccattaggcg aaggtaggat aatatagtaa atatcttatt tctcactgca gatgtaacaa 25740
gtagatgata attctctttt tgtcttgtca ggaagcattc taacaacttt tcaaaagatt 25800
tgtaatgtgc tctaaccctg ttgctccccc gccttgagtg ggagaagggt ggtcatttgc 25860
ttgtttgttt ccctagttta tcaaatatac aagatttcca cacacttctg tctttcatat 25920
tagaaaatat attacctaat gcttaccttt cattctgatt aagtttggtg ggcactgatg 25980
cttcaaaaat ttataaacat agtttgacaa ttcagtaaaa accatttata taaaatttta 26040
gcaaaatggc aattctataa gtcaaaataa gataaccagg taagattttt ttttttttt 26100
tgagtcggag tctcactctg ttgccaggct ggagtgcggt ggtgcgatct cggctcactg 26160
catcctccac ctcccgggtt caagtgattc tcctgcctca gcctcccaag tagctgggac 26220
tacaggtatg caccaccacg cccagctaat ttgtatattt ttagtagaga cgagatttca 26280
tcatgtggcc aggatggtct cgatctcttg acctcatgat ccacccgcct cagcctccga 26340
gagtgctgga attacagatg tgagccaccg tgcccagcca agatttttt ttttttagt 26400
aaaaactata tgtcaggccg ggtgtggtgg ctcatgcctg taatcccagc actttgggaa 26460
accaaggtag gcagattgcc tgagctcagg agttcaaaac caccctgggc aacgtggtga 26520
aaccctgtct ctactaaaat ccaaaaaaat tacctggcca tggtggcgca cacctgtctg 26580
tagtcccagc tactcgggaa gctgaggtgg gagaatcgct tgagcccagg aggcggaggt 26640
tgcagtgagc caagattgtg ccactgcact ccagcttggg caacagagtg agactccatc 26700
tcaaaaaaaa atatatatat atatatatat gtgtgtgtgt gtgtatatat gtatatgtgt 26760
atatgtatat atgtatacac acacacatct atctatctaa tttttaagtt gaaaatgaaa 26820
cttcagttct aggaaaggta tacggactct tcacttttca ctcttcacag ttctgtatta 26880
cattcattta ttttgtatgg taaaaattag cacttcagtt tttgccactt ttatcccatt 26940
cacttgatga tttgcgccag cattctagct taccatccag taccaagaaa ggaaatgtgc 27000
agggtatttt tgaaaaattg ccttattctg cacccatctc tagaggcaac ggaaacttta 27060
aattcttttc attgtcacag agagttacct ggagaaaaaa cagtctgctc agaccactta 27120
aattatccct catcgttgct tctctgtctc caaataccga aattctaaga aagcaagcaa 27180
aattattgtt aatgaaacat agtgtttatt ctttataata taggattgtt ctggatgtat 27240
tcattaaagt ggacattttg tgtaatatat atcctataag attgtgaggt aattctgttt 27300
tgaatgtaga gaattactgt gtcatctcta ctcctgtttt atattgttaa tcattagtcc 27360
cttgatcaag agttcagagt acaattttgt gggaaactgt aaccctgtag gacctctttt 27420
tatttcatga attccagaca taggaacact gtggccaaaa taacccaaaa atgcacgaaa 27480
gattcagcca cttcaaactg gctcccaggc tattttgatt ttatttgttt gtttttaatg 27540
tgttctggat attatccctt taacctaaat atatatgttt aaccttggca catacaggcc 27600
cacttaactt agtatctgtt tcttcttgga aatatcctaa ttttcaatta aatatattct 27660
tgatgtaatt tctcttttaa gatttgagaa acgaatttat attcccttgc cggaacccca 27720
tgcccgagca gcaatgttta aactgcacct agggaccact cagaacagtc tcacggaagc 27780
agactttcgg gaacttggga ggaaaacaga tggttattca ggggcagata taagtatcat 27840
tgtacgtgat gcccttatgc agcctgttag gaaagtacag tcagctactc attttaaaaa 27900
ggtaagtcat ttttatctat cctttcctat cttttttttt ttttttgag acagtctcac 27960
tctgttgccc aggctggagt gcagtggcac tatctgggct tgctgcaacc tccgcctccc 28020
gggttcaagc aattctcatg ccttagcctc ctgagtagct ggaattacag gcgtgtgcca 28080
ccacgcctgg ctaatttttt gtatttttag tagagacagg atttcaccat gttggccagg 28140
ctggtctcga actcctgact ccaggtgatc cacccacctc ggcctcccaa aggcgtgagc 28200
cactgcgcct ggccctttc ctgtattttt aaaatcaagt ttcaagaccg tcattggttt 28260
gtcttcaaag aatttgggag gatttcagtg ttgtataata ataccataag tgggtttaca 28320
taatagacat atttctgtgt cactgaaaca agttctttaa taaagtttat ttattggcca 28380
tttttggctg ggtgcagtgc cccatacctg taatcccagc actttaggag gccgaggtgg 28440
gcagatctct tgagctcagg agagaccagc ctcaggcaac atggcaaaac cccatctctg 28500
taaaaaatac aaaaattagc tggatgtggt ggcacactcc tgtagtccca gctacttggg 28560
aggctgaagt gggagaatct cttgagccca ggaggtggag gttgcagtga gccgagattg 28620
caccagtgta ctccagcctg ggtgacacag tgagaccctg tcttaaaaaa agaaaaggcc 28680
ttattaatca tcctaatgtg ttttaaattt tcctatgaaa tgatttttga agtatttgtg 28740
tattattttg atataagtat agaaacttaa atgtacagat gtaaggctgc ttctgacaca 28800
tctagaatat ttttgagtag aatctggtgc atttatacct cataaagtaa aaataataat 28860
agctaatatc ttttaaatat tcacatgcct ggcactgtgt tctgtctttt atcttactat 28920
ttcatttggt cttcacaaaa atcagcaaga gaggtacatg tagccagcca cctaatggtg 28980
gttagtcttg gacgtaggtg ctaacacagt agctagagtg ggacagacat gtgatttgcc 29040

tggaggtgag gtgtgtttcc aaatctgaaa tctcagccat tctgctatcc tgcccctcct 29100
gtgttctgtt ttgtctgcaa gaacaagtta agcatctgtg gctgagataa agttttataa 29160
gaaggctgtg gtcttgagat acccagtttg agtaacaatg aactaggtaa tctggaccaa 29220
ccctttttgct gtgaacaaca gtatccagta ttactcatga acatacatgt gaaaaactat 29280
tagcaaacct aatctcataa tgtatgtcaa gactaatatg actaagtcac atacgtatat 29340
ttatcatgta ccataataag atttttttaa gttaacaagc ttgaattaaa atctgaattt 29400
gaagcctggc accatttact tgttatacct ttttgagcaa attctattac ttctttaaag 29460
ctcatttta tcatttctaa aatggggatg tgataatgct attttgaagt cctttccaat 29520
aatgaagtga tataatactc ataaaacacc agtcgtggta cttcagatgc gtagggaatg 29580
ttactcatcc acccttcttc agccatttca ttactgaagt actgttaagt ttatgccaca 29640
gccattatgt tttttcttaa attctctttg ccatttagaa actataaatg cataggaact 29700
cagtcctttc ctactgagag tttgtcactc aattctgcat attactgcca gggtgagctc 29760
acatgacgct gaagccaggt ccttgttagt gggatgaagg tgacacagcc agatgagttt 29820
tatctctgca gctggggatc accagctgca gcagacctct cagcggctcc atcctttact 29880
tcttttaaag ctgattattt cagttaacaa gaagaagatt ccctgtttat tttttgcatt 29940
tactagtatt ttatttctaa aatactaatc tgggaccggt ggtagatcag aaagtagagc 30000
agtcaaaaaa gacaaaaatc cctgacctca tgcagctaaa atattagtga gtaaagacag 30060
taaaatacac ataataaggt gagagcatgt atttcttcac atctacacta atactgagat 30120
aatatttaaa gtcgattatt gtgaaacaga tgtgaaagtg cccccaagtt ggctgggcgc 30180
ggtggctcac gcctgtaatc ccagcacttt gggaggccaa ggtgggcgga tcacaaggtc 30240
aggggttcga gaccagcctg gccatcatga taaaacccca tctctactaa aaatacaaaa 30300
attagctcgg cgtcagggcg ggcgcctata atcccagcta ctcgggaggc tgaggcagga 30360
gaattgcttg aaccctggag gtggaagttg cagtgagctg acatcgcacc actgccctcc 30420
agcctgggct acaaagcgag actccgtctc aaaaaaaaaa aaaaagaagt gcctccaagt 30480
tgatgcagac tattgccatc taaatgttta catgtgacta tgtcacaaag attcttttttc 30540
tggatattgt taaaacttat aaataagttg tacatgataa aatggcattg atactgcatt 30600
tagcatacaa ttttgcactt agatttgaga gcatttttgt tgaaagagat tcagtagttt 30660
catactaaag ttccaaatat tacagaagtt tctcagctat agagcaactt ctgagaattt 30720
tttaatatat ctgtttttgg tctttaagat tactaccact tattaagcac tagaggggaa 30780
aaagagaaag ccacaggtaa ctatttttt tttaactcat cttctatgtg tgatcataca 30840
agcaatttat tgtaggaaat ttcaaaatat agaaagcatc aagaaaaaaa ctaaaagcag 30900
ctgggtgcag tcgctcacac ctgtaatccc agcactttgg gaggccgagg ctggcggatc 30960
acaaggtcag gagatcaaga ccatcttggc caacatggtg aaaccccgtc tctcctaaaa 31020
ataaaaaaaa attagctggg tgtggtggca catgcctgta atcctagcta ctcgggaggc 31080
tgagggatga gaatcccttg aacccaggag gtggaggttg cagtgagctg acatcgtgcc 31140
acttcactcc agcctggtga cagagcaaga cgccctttca aaaaaaaaac taaaagcaac 31200
caaaatctat cccattaccc aggggataac tacttttgt tgtatattct taatttttt 31260
tccaaaagta aacattctaa taaaacatac aaaactatta tacattttaa ctttcctggt 31320
aatggattaa gtcttaatgt tcaattttgt ttcaaataca tatttcagat tactgtttat 31380
ttcaggttcg cggaccttcc cgagctgatc ctaaccatct tgtagatgat ctgctaacac 31440
cttgctctcc aggtgaccct ggtgccattg aaatgacatg gatggatgtc cctggagata 31500
aacttttgga gccagttgtt tccatggttt gaaattttgt ttgtttttaa gaaaatattt 31560
taaaacatta tacttcataa ttttcagtta tataagattt gctggaaaat cttagtataa 31620
acaggtgtat tttacttaat attgttgact cattaaaaat tatgaagtta attgaactta 31680
aaatataaaa agtcattatt tttaaactat tttcttgggt ttagaggaac ctttaagtga 31740
aaaagctatt ttttaaagca caagattctg actttacagg tcttatagct agaatatctc 31800
caaaatgctc atatatagta taattattgt aaagacctaa ctaaaaaata gggttggtga 31860
aatatttttt ttcccctgca tctcagattt tttaaaataa ccagaagtca cctgtaatcc 31920
cagcactttg ggaggccaag gcgggcagac cacttgaggc caggagtttg aaactggcct 31980
ggccaacatg gtgaaatgcc atctctccta aaaatacaaa aattagccag gcagtgtggt 32040
gcacacctgc aatccagcta ctcagctgag gcaggagaat cacgtgaacc agggaggcag 32100
aggttgcagt gagccgagat tgcacaaccg caccccagcc tgggcaacag agcgagactc 32160
tgtctccaaa aaaaaaacaa aaaatagggc caggtacggt ggctcatgcc tgtaatccca 32220
acactttggg aggccgaggc gggtggatca tgaggtcagg agatcgagac catcctggcc 32280
aacacggtga aaccccatct ctactaaaaa tacaaaaaagt tagccagacg tggtggcacg 32340
tgcgtgtagt cccagctact tgggaggctg aggcaggaga atagcttgaa cccaggaggc 32400
ggaggttgca gtgagccgag atcataccac tgcactccag cctgggcaac agagcgagac 32460
tctgtctcaa aaaacaacaa caacaaaaac aaaaaccaga aatactaatg cttgtatcat 32520
tgtttttata gtgttttcta tttcacttct caaagctcca tagaaagttt aaatattgta 32580

gttctatttt tatttttaaa aattacagtt aatatagttt atattggaag aattaaaata 32640
atctacaaga cttactttgc aaggcatgtt atttttaggg aacttagggt gattgaaaag 32700
ttgtttactt acagaagaac attctcagtt ttcaaaattt tctgtcagtt tcagatcttc 32760
aaagaccttt aatggcaggt ttaatagtaa cttccagagg agaaactttt acacggtaca 32820
tagatataca tgtaaatgtg tatacaccca tatatttgtg tgtgaactaa ccataatgct 32880
tatctgagaa gttactcaga aacccctttg ttttaatcct tggtgtatta ttttttattt 32940
tttgtacttt gttcagcttt gaagaatttt tagtaactta gagtggtcct tttcattatt 33000
tttcatatcc ttttacagga gtctaggctt agaagttagt tttttaatga atctctgata 33060
ttagcatatt tttctaacca tttcgatctt aattgtaact aattatttta aaattctggt 33120
tagtagattt ttcatagaat ggtattctgc ttcacttaag tgtaagtgaa tgatattttc 33180
cttttgaatt atactgactt ggaaaattga tctcgtttaa gaaaaagaga taaggccaga 33240
aattagatca tgaacatact ttggaggcct tgatcctcag actaagggga ttgtccttct 33300
gataagtatt ggggagccat ttgatcagtt caggggaaga aagtatacgt ttctacacaa 33360
aggaataaca tgatcaaaga tcatgaaaga tctgtgaaag aataggggcc gggtgcaggg 33420
gcccacgcct gtaatcccag cacttcgcaa ggcagaggca ggcggatcac ctgaggtcag 33480
gggttcgaga ccagcctgac caacaaaacg aaaccccatc tctactaaaa atacaaaaat 33540
tagccaggtg tggtttcagg tgcctgtaat cccagctact tgggaggctg aggcaggaga 33600
atctcttgaa cctgggaggc agaggttgca atgagccaag atcacgccac tgcgctccag 33660
cctgagccaa cagagtgaga ctccaactca aaaaaaaaaa aaaagaatag ggagtttatc 33720
gttagtatct tatggcagac tgcaaatatt catatttaaa cttaggttaa agttcatagc 33780
aataatgata atatgacggt tttttatttt gttaactcta ttattaatat cctccctacc 33840
cttctttttc cctgacagtc ggatatgttg cggtcactat ctaacacaaa acctacagtc 33900
aatgaacatg acttgttgaa attaaagaag tttacagaag attttggtca agaaggctaa 33960
gccaaagaca aggaagatgc ttaccatatg tattctttct ttcatagata tttttgtcta 34020
tttggatcgc attaattgtt tccagtaaaa ctcttttacc acagggaaat acacatctca 34080
cttcagagtt ccattaggtt ttatattgta cttttcctcc attacttatt aaatactcct 34140
attaacaaaa ggtacaaaat aacaggttat gaggaaatga gcgatatatg aacggcataa 34200
aaacagaaat tacccagtaa aaaggatgtc agaaattgac atacaaatat ttacaatttt 34260
tatgaatggt ggtctttgca aagagcattt atattttctt tttttttac taaaatgata 34320
tagggtttat tttatatttt caaaaaaatt gttaaacatc attcttatca atgtaaaatt 34380
tacgttatta aaaaattaca aatgatagaa tctttactca gggatgggca ataaaacagc 34440
aaagagcttt gtgattgggt ttgaaagatt ttgaattata gacagtgctc ttaatagttt 34500
ttaataagtt gatatttttt tctgtgtaga tttaaataat ttctttaaaa gcgtaagctt 34560
tggctgggca tggtggctta tgcctgtaat cccagcactt tgggaggcta cgatgggtgg 34620
atcacctgag gtcaggagtt caagaccagt ctggccaaca tggtgaaatg ccgtctctac 34680
taaaaataca aaaattagct gggcatgatg gcgagtgtct gtaatcccag ctactcagga 34740
ggctgaggcg ggagaatcac ttgaacctgg gaggcagagg ttgcagtgag ctgagatcgc 34800
gctactgcat tccagcctgg gcgccgagat agcgccattg cactccagcc tgggcgactg 34860
aataaaactc tgtctcaaaa aaaaacaaaa acggtaagct tttatattta attaactgat 34920
acctttcact gtgcaatctc aagaggaaga ttttctaaat tagacattgt ctttacctta 34980
aagaagaaaa ttgtctattt cagtgtcttt tttagcaaga ttgaccaaaa caggtggtag 35040
gtgtgtaatt tttaactgtc atgaaatact gtaatgcgca cctttgctaa taaggaaatg 35100
gtctcccttt gaattatggg aagtttcttg tctgtgatat ggggtgtctg tgtacatata 35160
tgtatatttc ctttttttaat acaggataat aatatggggt cttgtccttc acctttaag 35220
ttcagggaac agtttgtcct gaccacacat atgttatcag tgtaatgctt tatgaagctt 35280
tagaaatgag ctcatgactc attaaaaaat aataaacgat gctattttaa tattttcaaa 35340
taaataatct gaggatgtgt ttcatacttt agcatcaacg ctgagttgtg ccttgtagaa 35400
ttgcattgtc tatacatgca aaagactatt aagttaggat gaactagatt ccctaattac 35460
agtgaaattc tatcttgcaa atgctataga ataacttgca ttttaaagtg tatttgcaca 35520
ttttacatat gctatgtggt tgcctttggg ttttctgtac agattgtttt tgttattaaa 35580
tggaaaaggc ctgaattatg ctcttatgtg aagtaaattg atataaccta ttgactgtat 35640
tctgtgtatt attctttatt ttgggacctt ctgtgtaaaa gaaaatatta cctaaattgg 35700
ttagatacaa agaaatcatt aaaattatta atgtatatat gaaaattctg tttttcaatg 35760
acttttatgt gaggattgag aactaactta gtaaactaac actgaattct ttgtggataa 35820
gataatttat acatcttacc tgtttatcag cctcagaatt atacagtact gctgtgtgtt 35880
ctttatggac agatactcaa aggaatccga atgggcatgt tctatttctc tggctcttga 35940
ccagtaatta taaagtttag gtagacattt tctgagccag caagtagatc atagttgtta 36000
agcatctcta tgacttacac cctggatatt tgaaagcatt ttgtctcacc ctgagaaatc 36060
tataaggact ctataattat aaaacatgtc aatagagaat tgacttctga gtccactgtc 36120

aactacatgc agaaaaacag aatgagtgta cttgattagt gccacagaag caccatttag 36180
ccactttcag agttgataaa gaggcatacc tacctgatgg tgctgtatgt tgcaagtgaa 36240
gcaacacagt gcccattgtt gacactaatg ttattttaag tttgtgagtt ttcacaggtg 36300
ccctttttctt tcctactttc atcccttcat ttgtatcctc ctcttatttt aatatggtat 36360
ctttcttttt tttgctagtc cacatcacgt atttccatta aaacctaatg atgatgaacc 36420
atgtaacctc ttgcaggaag accacccaaa ctacccccac cctacttttg ggtttcttta 36480
gcagttaaat tgctgctgct gctgtgttaa gatttttctc tggagcatgt tctattgatc 36540
atgtttagat catttagatt gtaagttctg agagttacat tttccccagc aattaatata 36600
gtgttgagtg cttttaggta cttagtaaag acttggggaa aaaaaaaggg agttatattt 36660
ttatttgat ctacctcatg gattatgtga gagacttata aaataggtaa gtttataaag 36720
tggctttgtt cttactggtc tagtgaagtt atgataaatt tctgtctcgg tgattgagaa 36780
ctacacaatc aaaaatgtat tcaaactagc tcaccctggc ccaccaaaac cgatagtaca 36840
tctttcccag ctgcgagttc agtgacctca tactagtaat ttaaagtctg ctctgatggg 36900
aatatttata ccatagaaac cagcaagagt tcagttagcc gtaaaaaatc aaggactgcc 36960
tgtaacttaa aataggaaga ttaatactat ttactgcatt tatggaaata aaggagacaa 37020
gttctttata gatgcataaa aagcaataca attcagcact tgctttttt ttaactgcac 37080
aaactaggaa tagaagggaa cttcttaata gaagaagttc tacaaaaaaa acctagagcc 37140
aacatattta attgtgaatt tttgaaaagc gtattcctcc agagatcagg aaaaaggcaa 37200
aggtgcccac tatcaccact ttttctcaac tttgtacaga ttctagccag tgcaagcaaa 37260
agaaatgaaa ggcttacaga taagaaataa aattgtcatt cacagatgac atgattgtag 37320
aaaattgata taaacgtaga aataatgtga atttaacaag gttgcaggat acaaagtcaa 37380
aattatctta gcaacaatta gaaattgaaa ttatgagagt ggtgtgtgcc tatagtccca 37440
actattgggg aattgaagtg ggaggatcac ttgagcccag gagttcgagg ctgcagtgag 37500
ctattaccac accactgcac tccagcctgg gtgagggaac gagaccctgt ctcaaaaaaa 37560
aaaaaaaaaa aaaggaaaat actaattaca atagtatcag agagcatcaa atgcctcaaa 37620

<210> 9
<211> 444
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:/note =
synthetic construct

<400> 9
Met Ser Ser Thr Ser Pro Asn Leu Gln Lys Ala Ile Asp Leu Ala Ser
1 5 10 15
Lys Ala Ala Gln Glu Asp Lys Ala Gly Asn Tyr Glu Glu Ala Leu Gln
20 25 30
Leu Tyr Gln His Ala Val Gln Tyr Phe Leu His Val Val Lys Tyr Glu
35 40 45
Ala Gln Gly Asp Lys Ala Lys Gln Ser Ile Arg Ala Lys Cys Thr Glu
50 55 60
Tyr Leu Asp Arg Ala Glu Lys Leu Lys Glu Tyr Leu Lys Asn Lys Glu
65 70 75 80
Lys Lys Ala Gln Lys Pro Val Lys Glu Gly Gln Pro Ser Pro Ala Asp
85 90 95
Glu Lys Gly Asn Asp Ser Asp Gly Glu Gly Glu Ser Asp Asp Pro Glu
100 105 110
Lys Lys Lys Leu Gln Asn Gln Leu Gln Gly Ala Ile Val Ile Glu Arg
115 120 125
Pro Asn Val Lys Trp Ser Asp Val Ala Gly Leu Glu Gly Ala Lys Glu
130 135 140
Ala Leu Lys Glu Ala Val Ile Leu Pro Ile Lys Phe Pro His Leu Phe
145 150 155 160

```
Thr Gly Lys Arg Thr Pro Trp Arg Gly Ile Leu Leu Phe Gly Pro Pro
                165                 170                 175
Gly Thr Gly Lys Ser Tyr Leu Ala Lys Ala Val Ala Thr Glu Ala Asn
            180                 185                 190
Asn Ser Thr Phe Phe Ser Ile Ser Ser Ser Asp Leu Val Ser Lys Trp
        195                 200                 205
Leu Gly Glu Ser Glu Lys Leu Val Lys Asn Leu Phe Gln Leu Ala Arg
    210                 215                 220
Glu Asn Lys Pro Ser Ile Ile Phe Ile Asp Glu Ile Asp Ser Leu Cys
225                 230                 235                 240
Gly Ser Arg Ser Glu Asn Glu Ser Glu Ala Ala Arg Arg Ile Lys Thr
                245                 250                 255
Glu Phe Leu Val Gln Met Gln Gly Val Gly Val Asp Asn Asp Gly Ile
            260                 265                 270
Leu Val Leu Gly Ala Thr Asn Ile Pro Trp Val Leu Asp Ser Ala Ile
        275                 280                 285
Arg Arg Arg Phe Glu Lys Arg Ile Tyr Ile Pro Leu Pro Glu Pro His
    290                 295                 300
Ala Arg Ala Ala Met Phe Lys Leu His Leu Gly Thr Thr Gln Asn Ser
305                 310                 315                 320
Leu Thr Glu Ala Asp Phe Arg Glu Leu Gly Arg Lys Thr Asp Gly Tyr
                325                 330                 335
Ser Gly Ala Asp Ile Ser Ile Ile Val Arg Asp Ala Leu Met Gln Pro
            340                 345                 350
Val Arg Lys Val Gln Ser Ala Thr His Phe Lys Lys Val Arg Gly Pro
        355                 360                 365
Ser Arg Ala Asp Pro Asn His Leu Val Asp Asp Leu Leu Thr Pro Cys
    370                 375                 380
Ser Pro Gly Asp Pro Gly Ala Ile Glu Met Thr Trp Met Asp Val Pro
385                 390                 395                 400
Gly Asp Lys Leu Leu Glu Pro Val Val Ser Met Ser Asp Met Leu Arg
                405                 410                 415
Ser Leu Ser Asn Thr Lys Pro Thr Val Asn Glu His Asp Leu Leu Lys
            420                 425                 430
Leu Lys Lys Phe Thr Glu Asp Phe Gly Gln Glu Gly
        435                 440
```

<210> 10
<211> 17315
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:/note =
synthetic construct

<400> 10

```
cctcccaggt tcaagcgatt ctcctgtctc agcctcccaa gtagctggga ttacaggtgc      60
ccaccaccat ggccggcgaa tattttgtat ttttagtaga gacagggttt caccatgttg     120
gccaggctgg tcttgaactc ctgacctcaa gtgatccacc tgcctcggcc tcccaaagta     180
ctgggattat aggcgtgagc caccgcgccc ggcctctcct ttgtatttaa ttagcatttt     240
gtgaagagag attaaggtgt gccaattttt tggttttttt ttttttttt ttttttttt      300
ttttttttt ttttggagg cagggtcttg cttcatttcc cagggtggag tacaatggcc      360
tggtcacggc tcactacagc ctcgacctcc caggctccag caattctccc aactcagctt     420
acggagtagc tgggactaca ggtgtgcgcc accatgccca gatacgtttt ttatttttg      480
tagcgatggg ggttcacctg ttgccgaggc tgttctcgaa ttcctcggcc caagcaatcg     540
gcctgcctcg gcttcccaaa gtgttgggat tacaggcatg agcccccata cccagcctca     600
acttggatgt tttttgccct ccatttttt tctctcacag gatagacttg tggacaaatg      660
gatcacttgc ctctggtgcg ttttttgtt tgtttgtttt gttttgtttc ttaatttagc      720
```

gacgaggtct tactatgttg cccaagctgg tcttgatcct cggctcaagc aatcctcctg 780
ccttggcctc ccaaagtgct ggaattacag gcgtgggcca gtacacccag ccctagaact 840
gctttgatca ttgtatcccc agaaactagc acacaacgct gcacacagta ggtgtgcatg 900
cctgtaatcc cagctactcg ggaggctgaa gcaggagaat cgcttgaacc cgggaggcgg 960
aggtcgcaat gagccgagat catgccactg cactccagcc taggcgacag agcaagactc 1020
cgtatcaatc aatcagtcaa tcaatccaag caaaacaaaa ctggctttga acagtcagcg 1080
gcatcccacc agggattgag ccccagtgtg gggcacactc tgtttacatt agctagctca 1140
tttattcccc acaaaagcca gacaaggcag gtactaatta tgatcgtgcg atattaaaca 1200
tggagaaacc aacgctctga taagttcgtt ttattgtgag aataaaactg ctcaatgggc 1260
gctgttattc caagcgaggt caggatttcc ggattctacc ccgctccgcc cctggccccg 1320
ccccgagaat gtacgtcccg ccaggccccg ccccggcctt gtcccaccca ccccgtcccc 1380
gccccctgtc cgtctcgccc tagtcccacc tcctcgtccc cagtcagtcc tgcctctcga 1440
accgcccccc gccttcgccg aggcagggag aggtgcctgc agtgcgagcc gtggacgcta 1500
gggggcagcc gctcggctcg ggcccgaccg ctctgggctg cgctcctcgc ttgccctcgg 1560
actcggctcc cgctgcgagc ggccgccctg cccgcgcacc gcgctcagcg cccaccgccg 1620
ggcttcccgc gccggaccca gtacctcggc tccccggggc cggaccgagg ccgcaagcag 1680
cgccgcgggg tgtggggcgg acccaggaga tgaaatgaca acgtcaaccc tccaggtact 1740
tcgtgcggcc cggcccgact tcggaggccg aaggcagcgg ggcctgcggg ccgcagagcc 1800
gggcggcggg cgggtggtca ggtgggcgcc tcccagcccc ggcctcgcga ccccgctccg 1860
gccgcccccc cgccgcttcc gcccgtgtct gtttccacgc gctcggagcc gggtggcatc 1920
tccaccctca ggccttccgg ccctgcccgg ctggcccagc cccagggccg cttcgtgtag 1980
ccggcgcccc ggtaccgccg tggggatcag cctcacccga gccgcacagg caccctcagg 2040
aagctcggcc cggggtccgg ctgggcgttc agacagcgat gctcgctccc tcctctcctg 2100
tctcgctccc tcctctcctg tcccgggtcc tggaagtccc tcgacctcac ttcgaacaag 2160
ccccggcccc ttccccaatc ttcgacttcc tcttctcggc cagcgaccct tcattcattc 2220
cctagatcct gcagggtggt gctgggatgg cgccacctgg gtgaacaact gacccaccct 2280
ctaggtgctg tcagtcctgc ggggccagaa gagatcaaca tgctcaggcc tcattccggg 2340
gtttgaatct cagctctgcc actcaccagc aagtgacctt gggctagttt ttttttttca 2400
tcttcgggtc tcgagtttcc tggtctgtaa aatggggtaa tgccctctag gttgttgtga 2460
tgatagcttt aaattgaata attacggtaa agcattcttg tatgtctggc acatagtagg 2520
tgtccattaa gtgacagcca acaaagtgcc tgggagagtt aggcaagact tcatggaaga 2580
agggacatca gagttccagt agcattcaga gacctggagg tggaggaggg cagaagtagg 2640
agcagacatg gaatgagttg gggtaccagg ttaggaatta agagtccata gccagatccc 2700
accgtaccac cagcagaagc cgcagatccc gtctgcccag gcaggttttt tatggcccgt 2760
gaaccgagaa tgtttttcag tttcaaacac aggtgtattt atttatttt aaatttttatt 2820
tatttatttt tttgagatgg agtctcactc tgtcacccag actggagtgc agtggtgtga 2880
tcttggctca ctgcagcctc tgcctctcag gttcaagcga ttctcctgcc tcagcctcct 2940
gagtagctgg gattacgggc acccgcaacc acgcccggct aatttttgta tttttagtag 3000
agaccgggtt tcaccgtgtt ggccaggctg gtctcgaact cccgacctca gatgatctgc 3060
ccacttcggc ctcccagagt gctgggatta caggtgtgag ccacgggtct ggcctaaaca 3120
cagctttatt aaggtataac ttgcgtgctt accatataac tcacccgtta tatatttttg 3180
tttttattaa ttttttttt tgagacagtc tcattctgtc acccaggctg gagtgcagtg 3240
gcacaatctc aactcattgc ggcctccgcc tcctgggttc aagtgattct cctgcctcag 3300
cctcccaagt agctgggatt acaggcacct cccatcacac ccagctaata ttttttttgta 3360
tttttagtag agacggggtt tcaccatgtt ggccaggctg gtttcgaact cctgacttca 3420
agtgtgatct gcccacctcg gccgcccaaa gtgctgggat tacagatgtg agccactgtg 3480
cccagcgcaa ctcacccatt taaagtgtat aattcagtgc ttttagtata ttcatggttg 3540
tgtgaccatc accataatca ctttttagaa tattttcatc acctccaaaa gaaacccctc 3600
acccacaacc cctctctctc ccagccctgg gcaaccactg ctaatctact ttttttctct 3660
gtgggtttgc ctacagagaa acccacttta tataaacaga gtcatacaat aggtggtcct 3720
ttgtggctag cttctttcat tatgacttaa cataatgtca taggttttta catttttaa 3780
atggcttgaa aaacagtatt tcatgacaca tgaaaatcct atgaaattca catttcaggg 3840
tctataaata aagttttatt ggaacacatc cctgttcgtt tattgatata ctgtctacag 3900
ctgcttttgt gccatcaagg tagagttgag tcgttgctac agagattatc tggctggcaa 3960
atcctaaatt atttaccatc tggtccttta cagtaagcat ttgccagccc ccattagcac 4020
tcttttcttt catcctcatc ttttgccata cgtgttacta gagatgcggc ggggctggca 4080
gtttctcacc ttaccctcct gttctgtaaa actgtttctc tccattgaaa cttggacccc 4140
ctgtgtgtga aaaactgtat gtggctgcca gaaaagtcca gtgcactctt ttttttttt 4200
tttttttga gatggactct tgctctggtg tgatctcggt tctctgcagc ctccacctcc 4260

tgggttcaag caattctcct gcctcagcct cccaagtagc cgtaactaca ggcacgcgtc 4320
accacacccg gctaactttt tgtattttta gtagagatgg tgcttcactg tgttagccag 4380
gatggtctca atctcctgac ctccggtgat ccgcccgcct ccgcgtccca aagtgctggg 4440
attacaggca tgagccaccg tgcccagccc cagcacactc ttaggtgtat tactagaatt 4500
tccaaaagga gggaggggat gggtcagtgc tgctccttat gggacaaaga ataccttcag 4560
tattgttcta aggatagacc caagcatgaa gctacagaga gtgaaaaagt ccaaggtact 4620
gggggtgtgt gttctggaga cgagaagcca ccagggagag gagcgctgcc ttcctgtctc 4680
atctgaagga acatcatgca gaagagggca gacaagggca gcctgtgcct aagctcattt 4740
ccaagaaggt ttggggctcc cagccattca ttgtaggtag tagagggagc agaggcttga 4800
aaataatggc cagcactcac ccctcttctc atctcttcaa tccgctgagt tcatttcaca 4860
tcagcctggc ccagcctccc ctggcctgtc tgttttgtcg agctgtctgt gtgtgtttct 4920
ttcggagact tcttccttgc agataagaga gggctgaaca aatgccccgg ccctgccaag 4980
agccggtcag gaggtttatt gtttgctgtg cctggaaccg tgggtctgcg gtgggactgc 5040
ctgcctggta actggtgccg gatctgagga ggagaattct ggtcatgctt cgtgcaggcg 5100
ttggtgaagc tgcggttttt aaaaccaggc tttaaaacca tgtaggtgct ttactgaccg 5160
gtcgcggtgg ctcatttctg taatcccagc actttggaag ccaaggcggg tggatcactt 5220
gatgtcagga gttcgagacc agcctggcca acatggtgaa actctgtctg cactaaaaat 5280
acaaaaatta gccggatgta gtggtgggca tctatcactt gaacccagga ggtggaggtt 5340
gcagtgagcc gagatcgcgc cagtgcactc cggcctagat gatagagtaa gactctgtct 5400
caaaaaataa taataataaa taaaaccacg taggtgcttt cagttcagcc taaggcactg 5460
gcctacagga tttatgagat tcccaaagac actcattctc acagagctct tagaaggttc 5520
cagagggaca cagttctctt ctgttggcag cagggagcaa gcaagtactt ttcttactgg 5580
agctactggt atgtggagca ttgtgtgtaa gcgtcccttc ccagggtgtc actttagaca 5640
gcgaaggaaa ctgtcttctc tctgcagcct ggagaagcaa gctgcatgga gccaggagga 5700
ggaggagaga ggcctttcag tgtatcctgt gcaaacttgc agtccgggca ggactttcct 5760
tctggctggc tcttctctgg ccctgaggcc gggtgcaatt tctggtctcc ttgctaaaca 5820
aaccagcagt cacttgctcc cgcgtgtccc ttctgccagg ctctcagctc gggtctgttt 5880
ctctttccct tgaggcagca gaatagaaag tcccatgaac ttgccccttg agccctgacc 5940
tgagatatta ggtcaggttt ccacttgctg aaatccagag cgtggctttt aaccgtctga 6000
gcttattgtg aagatggccg gctggttttc aaagcctcct ttgaaggcct ttctagccga 6060
tgttgaaaaa gtaggccaat gaaatggccc gcaagcaata aatccacagg cccagcggca 6120
tccctcgtca cgtttcatcc ccatgcctgt gaccccaggg actttccacc tctccgagga 6180
agcactgagc catttgcttt gcctttccag aaagccattg atctggtgac gaaagccaca 6240
gaggaggaca aagccaagaa ctacgaggag gcgctgcggc tgtaccagca tgcggtggag 6300
tacttcctcc acgctatcaa gtgtgagtca cacgaggggt cctcaggctg ccgcactcca 6360
gaggggagcg gaggagaggg ggtggcgcac gagcctcaca ggggcccctc tgtgttccct 6420
ttcacagatg aggcccacag cgacaaggcc aaggagagca ttcgagccaa gtgcgtgcag 6480
tacctagacc gggccgagaa gctgaaggat tatttacgaa gcaaagagaa acacggcaag 6540
aagccagtca aagagaacca gagtgagggc aaggggtgag tgtctgcagc gtccgcccag 6600
gaacctgggc cctcctcacg gctccctatc attcctggcg ctcatgctgc cttggacttc 6660
cctgtgggaa tggtcctcat agtgcaggga tgacagtgcc agcagctgct ttcagggagc 6720
atggcctgga gcacttccat gttcaccgtc tcattggatc tgacccacaa gccatgggga 6780
gaggggagc ttggtgcaga tggtagagcc tctcaggacg ggaccccagc tctgtcacct 6840
cccagcctag tgccaggacc ccgtgtgcct accatctcca catcttggaa gtggggaggt 6900
cgtggaactc acctccgagc gttgccacag ggagcagtga gatgacgtgg gcagtgccct 6960
tggagcttgg cacacgagag catgcaatga gtgccagctg ctgtgactgt caaagagtca 7020
cctgcctcca acaggtgggg cggccggctc gggttaggag ccctgccgag gtggcacagg 7080
atggaacttg gacccccaaa gcagggcggg tggccccgat tccccagcca gggtcctctc 7140
cacagcctct cacctcaggt actgggttcg cccaccacgg tgatcccagt catatgaaca 7200
caaatgctgg tggtatctcg gaggcgtggg aggggaagtg agggggtccc ccaggcagga 7260
gaggggtggg tggtgtcagc tgcctcacca gaggcgaggc ctcctcctgg atggagggtc 7320
tgcaccacgt tcccatccgc agagggagag gtctggcgag gcaggggtgg aggccaggga 7380
tctccaataa ggattaaatc aagttttgag gatgacagat tccatccagg gaggaagtag 7440
ccctggggtg acagtggctg cgttctcccc aggagttccc tttcgatttc ggttttgctt 7500
tctcctctgt ccttcggcag tggctgcttt ggagtcaggg gcacatgctg tccctgagtc 7560
tctacaaaaa gcttctggtg cggggcatgg tggctcacac ctgtaatccc agcactttgg 7620
gaggccaacg cgggcagatc acttgaggtc aggatttcaa aaccagcctg gccaacatgg 7680
tgaatcccgt ctctattaaa aatacaaaaa aataggccgg gcgcggtggc tcacgcctgt 7740
aatcccagca ctttgggaag ccgaggcggg cggatcacga ggtcaggaga ttgagaccat 7800

```
cctggctaac acggtgaaac cccgtctcca ctaaaaatac aaaatattag ccgggcgtgg   7860
tggcgggcac ctgtagtccc agctactcag gaggctgagg caggagaatg gcgtgaaccc   7920
gggaggcgga gcttgcagtg agccgagatc gagccactac actccagcct gggcgactga   7980
gtgcgactct gtcttaaaaa aaaaattagc caggcattgt ggcatgtgcc tgtaatccca   8040
gctactcggg aggctgacgc cagagaattc agagaattac ttgaacccgg gaggcggaga   8100
ttgcagtgag ctgagatgga gattgtgcca tcgcactcca acctgggcac catctcaaaa   8160
aaaaaaaaaa aaaaaaaaga agcttctggc agctgcattg tttgactggc ttttttttt   8220
ttttttcccc agacagggct tcactgtctc ccaggctgga gtgcagtggt gcgatcttgg   8280
ttaactgcaa cctctgcctc ccaagtagct ggcaccatag gtgcgtgcca ccatgcttgg   8340
ctaatttgt gttttttgta gagacaagat ctcgctgttt cccaggctgg tctcaaactc   8400
ctgggctcaa gcgatcctcc cgcctcagcc tcccagagtg ctgggattac aggggtgagc   8460
caccgcaccc agcctgggct gtgagagcca ctgcaggtgg cctcccagag tatgaggtgt   8520
gagtcaccgc acctggcctg ggctgctctt taagtccagc tgaagcacat ttcccttacg   8580
gccttcagga gagggccgac ccggattccc atctgctccg cgaagaaaca gggcgggtta   8640
taggagtgtt tgctgggggg gtcctgaagt gagctgctga gagcattggc ccaaatgtcc   8700
ctgaagtact tgcttgagtc acacggcact gttaagccag gaaggcttcg ttccttaacc   8760
agtgtgcagg cagcggagcc tggacttgct gggagcacct ctgtgctcca ggctgaaggg   8820
ccagcttgtg actttccgtc tcccttccca gcagtgacag tgacagtgaa ggggataatc   8880
cggagaaaaa gaaactgcaa gaacagctga tgggtaagtg gctcgcggcc ctgtggcagc   8940
ggcaggggat gagagtgggt ccgctgctgg gttggctcat gccccttggc cgggcccggg   9000
cccgggccgg cctccctctc gcaggtgccg tcgtgatgga gaagcccaac atacggtgga   9060
acgacgtggc cgggctggag ggggccaagg aggccctcaa agaagctgtc attttgccaa   9120
tcaaattccc acacttgttc acaggtgaga gttgagccat ttcaaacccc aatgtaaaaa   9180
ttccaggctt ccgcagggct ggcactccga ggcacccggg agtcagtggc gactcagccc   9240
ccggggcctg cagagggcca ggcctggctg ctcgctggtg tccacaacac tgctgtggct   9300
ggagctatga gatgacgatc acgtcacagg gtgggggcag gatcaggaag cagtgccagc   9360
cggtgtcact gtgcaaggtc atggggtcat agcaggggca ctgtgcaagg tcatggggtc   9420
atagcacagg gcacagcagg gaatgtagct cccatcactt gtttgccagt agagtccgtt   9480
gtttcacaga atgccctgtt ttactgtatg tatgggactc gagaccctct gcttttccta   9540
ttcctttccc cagtcaggag gcctaacttc tgtggttctc tgttgcaggc aagcgcaccc   9600
cctggcgggg gattctgctg ttcggacccc ctggcacagg gaaatcctac ctggccaaag   9660
ccgtggcaac agaggccaac aactccacct tcttctctgt gtcctcctca gatctgatgt   9720
ccaagtggct gggggagagt gaaaagtaag tcggccacca ggccatgctc tgccagcagc   9780
cccggcactg ctagtgaccc agcggcccac cctgtggagt cacctctcag aggagtggtt   9840
tggttttcag ggagggtacc caagagcagt gtgtcgctag cttatcggct gggacacagt   9900
cccaacaaag acaagtgaag gtggaaggga cctgggagag gtgaagtgct ccacaagaga   9960
agggaaagga agtgaggata tttttggggt ctgggttgcg gtccctgttg gggcagtgtg  10020
tgggggacat ggttccccac actgactcct gttttgcgtg tgatgagacg ggctcggaag  10080
aaggggactg acgtgttgta agtgcagagc tagcgcatgg ggctcctggg attgaggctc  10140
tgcgtttggg gctgacacca aggtgtgctc ctcctcttgg tgcagtgtgg cccgagggct  10200
cctcaccacc acgttttccg caattccggc atgaccgtgg cctgcgcctc cctgtgggaa  10260
gggtgagaag agggaagtgc cgggagccca ggcgggcacg gacgtgaacg tcttgtcctc  10320
acccccttc tcaccttcac aggctggtca agaacctgtt tgagctggcc aggcagcaca  10380
agccctccat catcttcatc gatgaggtgg attccctctg cgggtcccga aatgaaaatg  10440
agagtgaggc cgcccggagg atcaaaacgg agttcttggt ccagatgcag ggtgtgtgcc  10500
gggcccaggg ccccttggt tcttgttgca cctgaagcca accctgggct ttattctgag  10560
ctgcggctgg atttggttgt tctcagcctc ggagaggcac tccccagctc cagcccctca  10620
gggcacgggt ggacttcaat tcccaagagg tgcctgaggc ctctgcctca cgggccactc  10680
caccctccc atggcaggca gtgccatagg tctcacctgg cacagccaga ccaagatgtg  10740
gtttaatctc acttgggacc ctgccagtgg tgggtggcac agggatggct tcaattgctg  10800
acacacaaag ccccggggt ctgtccccag gtttcaactg acccgtgcag gtgtccagag  10860
tctgagtctt tgtctccctt tctccacagg ggtggggaat aacaatgatg ggactctggt  10920
tcttggagcc acaaacatcc catgggtgtt ggattcggcc atcaggagga ggtgagtctt  10980
ccccaggaga agccagggct ggaggcttcc tcccaccatg gtcacggtcc gcctgctgct  11040
ggcagcccgg gtgcagcctg gccctttttc cctggagtct tcccgtctgc ctgccaagca  11100
gagccctttg tgaggctggc ttgttgagga tgtgccgcca gcatcactgg cccatgaaat  11160
gcgtccgttt cactcaaatc ttcgtgcctc cccttccgtg aataccattc catcatccgc  11220
tgtcaactcc tgccgtgtgc tgggctctct aggtttgaaa aacgaattta tatccccttg  11280
ccggaggaag ctgcccgcgc ccagatgttc cggttgcatc tcgggagcac tccccacaac  11340
```

ctcacggatg caaacatcca cgagctggcc cggaagacgg aaggctactc gggcgcggac 11400
atcagcatca tcgtgcggga ctctctcatg cagcccgtga ggaaggtgca gtcggccaca 11460
cacttcaaaa aggtgagtgc ccgcggccac tgctgagaaa aatctcatag taagagcggg 11520
atgttcggtt ttttttttcc cagctcctgg tcctgctccc cggctgctca ggtgacacag 11580
tctctgaggc ctcctggaga gccgagggcc agtgccatgg gggctgcacc cactgtcccc 11640
tccctgcagc tcttctggag tgtggccatc tcagggtgtg tgaaagcaga ggacagggcc 11700
actttactgt cttaaccctg ctgcggcctc ctccgtcagc actgtgtgct cttgtgcggg 11760
gtggacacca aatcagtgtt tggaaagtgt tggatataaa atgaccagga agacctgtct 11820
attcattcag cagatctctg ctgagtattt gctgtgggtc ctcctgtgtg cccggccctg 11880
tcctaagtgc tgggaagata tgagtgacca acagacaaaa atcctagacc ttccgttgtt 11940
ggtggggaga cagtcaacac agtcagtgag tgtctcagag gatggagatg gagcagggtt 12000
aggagagagg tgtctcgggc atggagggaa gggcagatgg cactggggtg atgcctgaga 12060
caggaggtga aggtgtgagg gaatgagcct tgtggacatt gggggaatgt tttgccaagg 12120
cagcggactt cagttggaaa accactttgt gttaatctat tttcatgctg ctgataaaga 12180
catacccgag actgggaaga aaaagaggtt taattggcct tacagttcca catggctagg 12240
gaggcctcag aatcatggta ggagaccaaa ggcacttctt acatgctggt ggcaagagaa 12300
aatgaagaaa caaaagtgga aacccctgat aaacccatca gatctcgtga gacttcacca 12360
tcacgagaat agcacgggaa agaccagccc ccatgattca gttacctccc cctgggtccc 12420
tcccacaaca tgtgggaatt ctgggagata gatacagttc aagttgaaat gtgggtgggg 12480
gcacagctaa accatatcac acttttaatg tcactgtatt cctggtgaga cgtgacagtg 12540
ttaatgagga agacatcaag gcattggggc tacagattgc cagagtccca tgcaaatcat 12600
gagtcgctcg gaccacccag cccgagttct gaaggagccc gtcctcctca gatgctaggg 12660
acccacagag cattctcttt gggacttcct tagagaccgg aggggtcgag cccctctgac 12720
actgaggggc agctgcccca gtcagatcca tttggtttca ggtctgtggc ccctctcgca 12780
ccaaccccag catgatgatt gatgacctcc tgactccatg ctcaccaggg gacccaggag 12840
ccatggagat gacttggatg gatgtccctg gggacaaact cttagagcct gtggtttgca 12900
tggtaagtga cttgacaggg gagggaagag ggctgcacag agcccagaaa attgagggtt 12960
tgggtccaga ataaaaacgg tcttctccag gccaggcatg gtagctcacg cctgtaatcc 13020
cagcactttg ggaggccagg gtgggcagat cacaaggtta ggagttcgag accagcctag 13080
ccaatatgga gaaaccctgt ctctactaaa aatacaaaaa ttagccgggt gtggtggtgt 13140
gcacttgtag tcccagctac tcaggaggct gaggcagaag gatcgcttga agccgggagg 13200
cagaggttgc agtgagcaga gattgggcca ctgcactcca gcctgggtga cagagtgaga 13260
ctccgtctca aaaataaaaa atagtcttct tctatgtcct gggctgttga gaatggcctt 13320
tgtagctcag ggaagctggc attacactca ctgggggccc agaatgcagg tggcagcatt 13380
agcattttgt acatacagag aaaggctaga tgcagtggtg cagtgatagc tcactgcagg 13440
cttgacctca cgagctcaca caatcttctg acctcagcct cctgagtagc aacaggcaca 13500
caccatcatg cccagctaat ttttcctttt ttttttagaa acggcctcac agtgttgccc 13560
aggccagtct gaaactgctg ggctcaagtg atcctctcag cgtggcctcc caaagtgctg 13620
ggattacaag catgagctga gccaccacac ctagcctgca ttttatattc tttagcaagt 13680
aaagtcaggt ccagataata tgacgtcaat ccatgggttt ggctttgcgg tgcagtgagg 13740
cagtgaatgt gtcaagactg tgacgcagtt gcaaaggcag cgcctcacct gcatagcctg 13800
gcccgtgcca tcctggcaag ggtctgccct tgtggctggt gttgggggtg tgcagctaag 13860
aaggctcctg gccgggcgcc gtggctcccg cctgtaatcc cagcacttcg ggaggccgag 13920
gcgggcagat cgcctgaggt caggagttca agagcagcct ggccaacata gtgaaacccc 13980
gtctctgcta aaaatacaaa aatcagccat gtgtggtggc acgtgcctgt agtcccagct 14040
atttgggagg ctgaggcagg agagtcactt gaaccgagga ggtggaggtt gcggtgagcc 14100
gagatcatgc cactgcactc cagcctgggt gacagactcc gtctccaaaa aaaaaaaaaa 14160
aaaaaaagaa agcaagttcg taaggggagt tatgagatga tggagcttag attctcagac 14220
caggaggagg cccatctgag ctgcctcgag agggaagagg gtccctgcca taggcagatt 14280
caagcaggcc tctggggtgg ccttgaaggc aaacctcttc ccaccctcag ctgcgcctgt 14340
tgtgtgctga ggggaggttg gggacctgga gcctggctcc tgctcaggca catactgttg 14400
ccttcacagt cggacatgct gcggtctctg gccaccaccc ggcccacggt gaatgcagac 14460
gacctcctga aagtgaagaa attctcagag gactttgggc aagagagtta aaagctgctt 14520
cacttgggca atggtgaagg tgggaggttg attggggcaa atccaggcac tccccatgtc 14580
aacagccaga cagggctcca gggcttgtcc cagtcaatac agagttccct ctgctgtctg 14640
gccgtctgcc agggagccag aaggaagggc cttgcagcca cagagacact ccactgccct 14700
ggggcacaca gtggacactg ctcttcctac ttcctcctct cctggatgct catcagctcc 14760
ttctgcctcc ccccctttt tttccatctt ttgttccct aaattaatgc tgcttggatt 14820
ttcatcttat ttataaagat aaaatcacct ggaagtgtca aggagtgggg cggggtggcg 14880

```
ggggagaagc agccgtgctg ccaggtcacc cagacctcca gacagccggc tagccccact   14940
gcccgttcct tttacgccca agttttgctc cttgagagca gattggctga tgcccctgca   15000
accccagccc aagctctgcc tcaaagaccg agtgacataa gccattccca ccctcctagg   15060
ttcacatcca gggctgtgtc ttccttgggg gaggagatgg tgtcgtttag atcagggtaa   15120
ggcagtcagg cgggtgttca ccactgcctt ttcttcctct gagcgtgaga acactgaacc   15180
cagccactgc ccctgggtcc ctgtcctgga aatggtctaa taaatccttt tcccttcttg   15240
agctacccaa gtgatctcat ctttccttga ctcttaactg agactctgaa gggccaccct   15300
tgcttccaat taaaagaccc tctggttttc tgtctcctct cttttctaaa gcagagcagg   15360
aaacagaaca ctgcagttct taggcctgtt ctccagtcag gtgtgcaagg cctcctcccc   15420
ttaaggcctc agaagtttgg ctggggatgt ttgtgggatc cagacagttc ttgccgttgt   15480
tcggcctaca gatcagagac tgcaaagtgg gagccctgcg taccagaccc gtcctgcagc   15540
tgtgtttgtt tggtttggtc tgccgctaac atttaaaagt cgagagttgc tgggcgcggt   15600
ggctcacgcc tgtaatccca gcactttggg ccgctaacat ttaaaagtcg agagttgctg   15660
ggcgcggtgg ctcacgcctg taatcccagc actttgggag gccgacgcgg gcggatcaca   15720
aggtcaggaa atcgagaaca tcctggctaa cacagtgaaa ccccgtctca ctaaaaataa   15780
aaaaattagc cgggcgtggt ggtgggcgcc tgcagtccca gctacttggg aggctgaggc   15840
aggagaatgg cgtgaacccg ggggacggag cttgcagtga gccgagatcg cgccactgca   15900
ctccagcctg ggcgacagag caagactctg tctcaaaaaa aaaaaaaaag tcgagagttg   15960
acataaaaat tcagaattct gacttctaga agattaaatt aggtaatgtt gggcctgcgt   16020
cctcatgggg cagcctgtca gcggccacta agtagaggct gctctgtaca gaaggggtcc   16080
cccatcccca ctcatctgtg ttgcctgccc agatagcccc atggccattt cagtttgcga   16140
cccccaacag accaaggacc cagagctggc agccctggac ccacgtccgg caggcagtga   16200
agctgcagag gctgtcccac gtttacccca tgcatataaa cagaccaaag tcaaagcaca   16260
ttcacacaga aagacctggc ctagtaaatg aagcttatag cctgcaaagg gctttgttcc   16320
aaccttccct ctccaaaacc cagtactgct cctctacctt tcaaacctgg ttgggccaga   16380
actgattaca tggccaactc ttggctcaca ctgggcttgc actgggccat ttctttgctc   16440
tggttcagag cctctgcaga agtgagctaa ctttggatca gtagctccaa cacttggctc   16500
cagtgctggc aggttacccc tcaaccatag cagctgggat ctttgtggct tgggggtgct   16560
tgtgtaagtt cttggttatg tgccctcagc caccaggata tgaaccgtgt ctgcagagct   16620
tcactttagt gagcttacaa gtttttaaat ttttacttct agttacattt acttgaatca   16680
gaacgttgtt tcctcattcc tactgcttaa acaccttgac aagtcctagg ggttaacaaa   16740
ggttagcatg gctatggtcc atccctgtgc tctagttaga gcgtgaagac acctgacttt   16800
ccagttgtct ctctccatga ccagcaacag gaaccactga cgctgaactt tggacagtgg   16860
cctcagactc tggctgccag cacacaacct gccatcatcg atgttaaaca tgctgacatg   16920
tgcagaggag tttcctccct gaaatgctct gaaattcact tctctgcctg gggattctgt   16980
tataaacctt ctgcctacat tggctttcac tgtggaagtt gatttctaaa actctgatga   17040
gctcacaacg ttgacatgta tatgcgtttt ttgtgagtgc ttcctgctca aagtgggaga   17100
gattttactg gaaatgcaat aaagtttgca ttttattgct accgatgccc accacaaaag   17160
cgttaacctc agaaagctat aagaagcatc aaaagattgc catagccccc tgtaagagct   17220
tctgaacatt gtccaaatga aaggaggctc agctgaattc taattttcct aaagacccat   17280
gtttttagca tctatgagct ttctacaaat ctccc                              17315
```

<210> 11
<211> 437
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:/note =
synthetic construct

<400> 11

```
Met Thr Thr Ser Thr Leu Gln Lys Ala Ile Asp Leu Val Thr Lys Ala
 1               5                  10                  15
Thr Glu Glu Asp Lys Ala Lys Asn Tyr Glu Glu Ala Leu Arg Leu Tyr
            20                  25                  30
```

```
Gln His Ala Val Glu Tyr Phe Leu His Ala Ile Lys Tyr Glu Ala His
        35                  40                  45
Ser Asp Lys Ala Lys Glu Ser Ile Arg Ala Lys Cys Val Gln Tyr Leu
    50                  55                  60
Asp Arg Ala Glu Lys Leu Lys Asp Tyr Leu Arg Ser Lys Glu Lys His
65                  70                  75                  80
Gly Lys Lys Pro Val Lys Glu Asn Gln Ser Glu Gly Lys Gly Ser Asp
                85                  90                  95
Ser Asp Ser Glu Gly Asp Asn Pro Glu Lys Lys Lys Leu Gln Glu Gln
            100                 105                 110
Leu Met Gly Ala Val Val Met Glu Lys Pro Asn Ile Arg Trp Asn Asp
        115                 120                 125
Val Ala Gly Leu Glu Gly Ala Lys Glu Ala Leu Lys Glu Ala Val Ile
    130                 135                 140
Leu Pro Ile Lys Phe Pro His Leu Phe Thr Gly Lys Arg Thr Pro Trp
145                 150                 155                 160
Arg Gly Ile Leu Leu Phe Gly Pro Pro Gly Thr Gly Lys Ser Tyr Leu
                165                 170                 175
Ala Lys Ala Val Ala Thr Glu Ala Asn Asn Ser Thr Phe Phe Ser Val
            180                 185                 190
Ser Ser Ser Asp Leu Met Ser Lys Trp Leu Gly Glu Ser Glu Lys Leu
        195                 200                 205
Val Lys Asn Leu Phe Glu Leu Ala Arg Gln His Lys Pro Ser Ile Ile
    210                 215                 220
Phe Ile Asp Glu Val Asp Ser Leu Cys Gly Ser Arg Asn Glu Asn Glu
225                 230                 235                 240
Ser Glu Ala Ala Arg Arg Ile Lys Thr Glu Phe Leu Val Gln Met Gln
                245                 250                 255
Gly Val Gly Asn Asn Asn Asp Gly Thr Leu Val Leu Gly Ala Thr Asn
            260                 265                 270
Ile Pro Trp Val Leu Asp Ser Ala Ile Arg Arg Arg Phe Glu Lys Arg
        275                 280                 285
Ile Tyr Ile Pro Leu Pro Glu Glu Ala Ala Arg Ala Gln Met Phe Arg
    290                 295                 300
Leu His Leu Gly Ser Thr Pro His Asn Leu Thr Asp Ala Asn Ile His
305                 310                 315                 320
Glu Leu Ala Arg Lys Thr Glu Gly Tyr Ser Gly Ala Asp Ile Ser Ile
                325                 330                 335
Ile Val Arg Asp Ser Leu Met Gln Pro Val Arg Lys Val Gln Ser Ala
            340                 345                 350
Thr His Phe Lys Lys Val Cys Gly Pro Ser Arg Thr Asn Pro Ser Met
        355                 360                 365
Met Ile Asp Asp Leu Leu Thr Pro Cys Ser Pro Gly Asp Pro Gly Ala
    370                 375                 380
Met Glu Met Thr Trp Met Asp Val Pro Gly Asp Lys Leu Leu Glu Pro
385                 390                 395                 400
Val Val Cys Met Ser Asp Met Leu Arg Ser Leu Ala Thr Thr Arg Pro
                405                 410                 415
Thr Val Asn Ala Asp Asp Leu Leu Lys Val Lys Lys Phe Ser Glu Asp
            420                 425                 430
Phe Gly Gln Glu Ser
        435
```

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
13 May 2004 (13.05.2004)

PCT

(10) International Publication Number
WO 2004/039311 A3

(51) International Patent Classification[7]: A61K 31/105, 31/095

(21) International Application Number: PCT/US2003/015722

(22) International Filing Date: 15 May 2003 (15.05.2003)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/381,027	15 May 2002 (15.05.2002)	US
60/382,702	21 May 2002 (21.05.2002)	US
60/451,146	28 February 2003 (28.02.2003)	US

(71) Applicant *(for all designated States except US)*: **UNIVERSITY OF UTAH RESEARCH FOUNDATION** [US/US]; 615 Arapeen Drive, Suite 110, Salt Lake City, UT 84108 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **SUNDQUIST, Wesley, I.** [US/US]; 299 Fairfax Circle, Salt Lake City, UT 84103 (US). **GARRUS, Jennifer, E.** [US/US]; 2686 Preston Street, Salt Lake City, UT 84105 (US). **VON SCHWEDLER, Uta, K.** [US/US]; 676 S. 1200 East, Salt Lake City, UT 84102 (US).

(74) Agents: **HUIZENGA, David, E.** et al.; Needle & Rosenberg, P.C., The Candler Building, 127 Peachtree Street, Atlanta, GA 30303-1811 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *with international search report*

— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments*

(88) Date of publication of the international search report: 14 October 2004

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: HUMAN VPS4A AND VPS4B FUNCTIONS IN VIRAL BUDDING

(57) Abstract: Disclosed are compositions and methods related to retroviral budding.

WO 2004/039311 A3

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US03/15722

A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) : A61K 31/105, 31/095
US CL : 514/707, 706

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
U.S. : 514/707, 705

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
DIALOG: EMBASE, BIOSIS, MEDLINE, CA SEARCH; WEST

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US5,981,602A (TATARINTSEV et al) 09 November 1999 (9.11.1999), Abstract, Claims 1-5.	1-39

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search
28 June 2004 (28.06.2004)

Date of mailing of the international search report
05 AUG 2004

Name and mailing address of the ISA/US
Mail Stop PCT, Attn: ISA/US
Commissioner for Patents
P.O. Box 1450
Alexandria, Virginia 22313-1450
Facsimile No. (703) 305-3230

Authorized officer
James Housel
Janice Ford
for

Telephone No. 703-308-1235

Form PCT/ISA/210 (second sheet) (July 1998)